

ONLINE SEARCH REQUEST FORM

USER SUSAN LORING SERIAL NUMBER 08/249,182ART UNIT 1806 PHONE 308-3998 DATE 9/21/94

Please give a detailed statement of requirements. Describe as specifically as possible the subject matter to be searched. Define any terms that may have special meaning. Give examples or relevant citations, authors, or keywords, if known.

You may include a copy of the broadest and or relevant claim(s).

Please:
search SEQ ID 1-11 and autotaxin (ATX)
autocrine motility factor (AMF)
motility stimulating autocrine
and antibodies which bind

[ATX does NOT MEAN ADULT THYMOMIZED
IN THIS APPLICATION]

OK

Susan, there are two (2) parts
to the 1G search. It is
as follows:
1) Pt. 1-2 - The key terms
are searched in all
standard AA data banks
then, comb. ID's searched
against Seq. ID's.
2) Pt. 2-2 - Seq. ID's individ.
searched.

Beverly
© 4994
12C14

RECEIVED
PTO
SEP 21 1994
TOLSON

08/249,182
Dialog
9/21/94
SKC

b 411

{>>Invalid Option Number

*** DIALOG HOMEBASE(SM) Main Menu ***

Information:

1. Announcements (new files, free connect time, price changes, etc.)
2. Database, Rates, & Command Descriptions
3. Help in Choosing Databases for Your Topic
4. Customer Services (telephone assistance, training, seminars, etc.)
5. Product Descriptions

Connections:

6. DIALOG Menus(SM)
7. DIALOG Business Connection(R), Headlines(SM), Medical Connection(SM)
8. DIALOG SourceOne(SM) Document Delivery
9. Data-Star
10. Other Online Menu Services & Files (MoneyCenter(R), OAG, TNT, etc.)

/H = Help

/L = Logoff

/NOMENU = Command Mode

Enter an option number to view information or to connect to an online service. Enter a BEGIN command plus a file number to search a database (e.g., B1 for ERIC).

?b411

>>Invalid Option Number

*** DIALOG HOMEBASE(SM) Main Menu ***

Information:

1. Announcements (new files, free connect time, price changes, etc.)
2. Database, Rates, & Command Descriptions
3. Help in Choosing Databases for Your Topic
4. Customer Services (telephone assistance, training, seminars, etc.)
5. Product Descriptions

Connections:

6. DIALOG Menus(SM)
7. DIALOG Business Connection(R), Headlines(SM), Medical Connection(SM)
8. DIALOG SourceOne(SM) Document Delivery
9. Data-Star
10. Other Online Menu Services & Files (MoneyCenter(R), OAG, TNT, etc.)

/H = Help

/L = Logoff

/NOMENU = Command Mode

Enter an option number to view information or to connect to an online service. Enter a BEGIN command plus a file number to search a database (e.g., B1 for ERIC).

?b 411

SYSTEM:HOME

21sep94 07:07:43 User214323 Session D226.2

\$0.03 0.002 Hrs FileHomeBase

\$0.03 Estimated cost FileHomeBase

\$0.02 TYMNET

\$0.05 Estimated cost this search

\$0.31 Estimated total session cost 0.013 Hrs.

File 411:DIALINDEX(R)

DIALINDEX(R)

(c) 1994 Dialog Info.Svcs.

*** DIALINDEX search results display in an abbreviated ***

*** format unless you enter the SET DETAIL ON command. ***

?sf all

You have 335 files in your file list.

(To see banners, use SHOW FILES command)

?s autocrine(w)motility(w)factor?(or autotaxin or ATX{

Your SELECT statement is:

s autocrine(w)motility(w)factor?. or autotaxin or ATX.{

Items File

```
{          5      5: BIOSIS PREVIEWS(R)_1969-1994/OCT W2
          1      6: NTIS_1964-1994/Oct B2
{({({      Examed  50 files
          2      73: EMBASE_1974-1994/ISS 36
          1      76: Life Sciences Collection_1978-1994/Aug
{          2      144: Pascal_1973-1994/Aug
          3      155: MEDLINE(R)_1966-1994/Nov W2
      Examed 100 files
          1      156: TOXLINE(R)_1965-1994/Sep
          3      159: CANCERLIT(R)_1963-1994/Sep
{      Examed 150 files
          1      265: Fed. Res. in Progress_1994/Aug
          1      266: Fed. Res. in Progress_1994/Aug
{          1      345: INPADOC/Fam.& Legal Stat._1994/UD=9436
          1      351: DERWENT WPI_1981-1994/UD=9430;UA=9426UM=9417
          1      357: Derwent Biotechnology Abs_1982-1994/Oct B1
          1      358: Current Biotech Abs_1983-1994/Oct
          19      398: CHEMSEARCH(TM)_1957-1994/Aug 13341571RNS
          3      399: CA Search(R)_1967-1994/UD=12110
      Examed 200 files
{          4      434: Scisearch(R)_1974-1994/Aug W4
{          3      440: Current Contents Search(R)_1990-1994/Aug W2
      Examed 250 files
          1      669: Fed.Register_1988-1994/Sep 20
      Examed 300 files
```

19 files have one or more items; file list includes 335 files.

?sf{({({({

>>>Unrecognizable command.

?rf

Your last SELECT statement was:

S AUTOCRINE(W)MOTILITY(W)FACTOR?. OR AUTOTAXIN OR ATX.{

Ref	Items	File
---	-----	----
N1	19	398: CHEMSEARCH(TM)_1957-1994/Aug 13341571RNS
N2	5	5: BIOSIS PREVIEWS(R)_1969-1994/OCT W2
N3	4	434: Scisearch(R)_1974-1994/Aug W4
N4	3	155: MEDLINE(R)_1966-1994/Nov W2
N5	3	159: CANCERLIT(R)_1963-1994/Sep
N6	3	399: CA Search(R)_1967-1994/UD=12110
N7	3	440: Current Contents Search(R)_1990-1994/Aug W2
N8	2	73: EMBASE_1974-1994/ISS 36
N9	2	144: Pascal_1973-1994/Aug
N10	1	6: NTIS_1964-1994/Oct B2

19 files have one or more items; file list includes 335 files.

- Enter P or PAGE for more -

E10 0 1 (((CARBOXYMETHYL)IMINO)BIS(ETHYLENENITRIL))TE
 E11 0 1 (((IODOACETYL)AMINO)ETHYL) 5 NAPHTHYLAMINE 1 S
 E12 0 1 (((PARA CUMENYLCARBAMOYL)METHYL)IMINO)DIACETIC

Enter P or PAGE for more

?exs

Executing TD092

24783 AUTOCRINE

143245 MOTILITY

0 FACTOR?.

0 AUTOCRINE(W)MOTILITY(W)FACTOR?.

46 AUTOTAXIN

0 ATX.{

S1 46 AUTOCRINE(W)MOTILITY(W)FACTOR?. OR AUTOTAXIN OR ATX.{

?6 ATX

S2 613 ATX

?s s1 or s2

46 S1

613 S2

S3 649 S1 OR S2

?s antibod8

S4 0 ANTIBO.D?

?6santibod?

{{Processing

Processed 10 of 11 files ...

Completed processing all files

S5 1718826 ANTIBOD?

?s s5 and s3

>>>Unrecognizable Command

?s s5 and s3

1718826 S5

649 S3

S6 73 S5 AND S3

?rd

>>>Duplicate detection is not supported for File 398.

>>>Records from unsupported files will be retained in the RD set.

>>>Record 440:4585345 ignored; incomplete bibliographic data, not retained in RD set

>>>Record 440:3418381 ignored; incomplete bibliographic data, not retained in RD set

...examined 50 records (50)

...completed examining records

S7 47 RD (unique items)

?t s7/7/1-47

7/7/1 (Item 1 from file: 5)

DIALOG(R)File S:BIOSIS PREVIEWS(R)

(c) 1994 BIOSIS. All rts. reserv.

7060647 BIOSIS Number: 87121168

SPECIFICITY OF ANTIBODIES TO SEA ANEMONE TOXIN III AND IMMUNOGENICITY OF THE PHARMACOLOGICAL SITE OF ANEMONE AND SCORPION TOXINS

BAHRAOUI E M; EL AYE M; GRANIER C; BERESS L; ROCHAT H

LABORATOIRE DE BIOCHIMIE, FACULTE DE MEDECINE SECTEUR NORD, BOULEVARD PIERRE DRAMARD, F-13326 MARSEILLE CEDEX 15, FRANCE.

EUR J BIOCHEM 180 (1). 1989. 55-60. CODEN: EJBCA

Full Journal Title: European Journal of Biochemistry

Language: ENGLISH

Toxin III (ATX III) of the sea anemone (Anemonia sulcata) is a polypeptide containing 27 amino acid residues. It has no sequence similarity with other toxins (ATX I and II) from the species, or with scorpion toxins, although they apparently act in a similar manner by prolonging action potentials. The specificity of ATX III antibodies was characterized using ATX III, ATX I, native and chemically modified ATX II, and scorpion .alpha.-toxins. The results obtained suggest that a region of

ATX III, partially or totally overlapping the pharmacological site shared with ATX I and ATX II, is immunogenic. It includes a guanidino and at least two carboxylate groups. The corresponding region is not immunogenic in ATX I and ATX II. Anti-(ATX III) antibodies recognize the similar regions of ATXI and ATX II and apparently do not recognize scorpion toxins.

7/7/2 (Item 2 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

6478959 BIOSIS Number: 85079480

CELLULAR PATHWAYS FOR REJECTION OF CLASS-I-MHC-DISPARATE SKIN AND TUMOR ALLOGRAFTS

SMITH D M; STUART F P; WEMHOFF G A; QUINTANS J; FITCH F W
UNIV. CHICAGO, COMMITTEE ON IMMUNOLOGY, DEP. PATHOLOGY, BOX 414, 5841, S. MARYLAND AVE., CHICAGO, ILLINOIS 60637.

TRANSPLANTATION (BALTIMORE) 45 (1). 1988. 168-175. CODEN: TRPLA

Full Journal Title: TRANSPLANTATION (Baltimore)

Language: ENGLISH

We have investigated the relative roles of the Lyt-2+ and L3T4+ T lymphocyte subsets in rejection of class-I-MHC-antigen-disparate skin and tumor allografts. To deplete T cells in vivo, rat anti-Ly-2 or anti-L3T4 monoclonal antibodies (mAb) were administered to adult-thymectomized (ATX) recipient mice prior to transplantation. BALB/c (H-2d) recipient mice injected the Ia-Sarcoma I (Sa1) (H-2a) tissue culture-derived tumor after depletion of the L3T4+ T cell subset in vivo. In contrast, depletion of the Lyt-2+ T cell subset permitted lethal tumor growth in all recipient mice. To determine the role of particular T cell subsets in rejection of Ld class-I-MHC-antigen-disparate allografts, BALB/c skin was transplanted to BALB/c-H-2dm2 recipient mice. Skin grafts were rejected by control mice with a mean survival time (MST) of 14.5 days. The MST of skin grafts for mice treated with anti-L3T4 mAb was 16.6 days. In contrast, administration of anti-Lyt-2 mAb alone (MST = >47 days) or together with anti-L3T4 mAb (MST = >50 days) caused prolonged or indefinite graft survival in all recipient mice. Depletion of specific T cell subsets was confirmed by flow cytometric analysis and by analysis of T cell function in vitro. These results suggest that Lyt-2+ T lymphocytes are essential for rejection of class-I-MHC-disparate allografts; indirect presentation of alloantigen to L3T4+T cells may not be necessary for rejection.

7/7/3 (Item 3 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

5418734 BIOSIS Number: 82063537

FUNCTIONAL CHARACTERIZATION OF PRETHYMIC T CELLS COMMITTED TO ALLOREACTIVITY

BLANK M; GRONAU C; SAUER M; WOTTGE H-U; MUELLER-RUCHHOLTZ W
DEP. IMMUNOL., UNIV. KIEL, BRUNSWIKERSTR. 2-6, D-2300 KIEL, FRG.
TRANSPLANTATION (BALTIMORE) 41 (6). 1986. 759-765. CODEN: TRPLA

Full Journal Title: TRANSPLANTATION (Baltimore)

Language: ENGLISH

The results of previous experiments on MHC fully allogeneic bone marrow transplantation (BMT) in non-thymectomized recipients indicated that anti-MHC alloreactivity starts to become irreversibly committed at the prethymic level. This is a matter of some controversy. Since it is possible that conflicting results depend on the methods chosen, we reexamined our previous results by applying two new approaches. Adult thymectomized (ATX) Balb/c mice received a syngeneic fetal thymus either 3 weeks before or 3 weeks after lethal irradiation and reconstitution with C57BL/6 BM incubated in antiserum. Since monoclonal antibodies such as anti-Thy 1 are of limited value for investigations of the above type (Thy 1 antigen crosses the prethymic/thymic border), we used two highly selective, excessively cytotoxic xenoantisera for incubation of the donor BM-either a specific

anti-T cell serum (SAT) that eliminated only mature T cells, or a specific antilymphocyte serum (SAL) that reacted additionally with a subset of prethymic T cells (PTC). In both experimental approaches the results were similar: (1) Recipients of SAT-BM developed antihost reactivity, in contrast to recipients of SAL-BM. (2) SAT-BM recipients became immunodeficient, whereas SAL-BM chimeras were immunocompetent. (3) Late mortality was observed only following SAT treatment. (4) Preliminary morphological findings in the lymphoid tissues of BM recipients agree fully with the functional observations. We conclude that the data confirm our previous results in nonthymectomized BM recipients-i.e., PTCs initiate antihost reactivity in MHC fully allogeneic BMT-and PTC commitment is thymus/thymus factor independent. These conclusions are discussed with regard to the problems of MHC allogeneic clinical BMT.

7/7/4 (Item 4 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

4437760 BIOSIS Number: 78011583

INJECTION OF MOUSE THYRO GLOBULIN AND-OR ADULT THYMECTOMY DO NOT BREAK TOLERANCE TO THYRO GLOBULIN DURING THE LUPUS-LIKE GRAFT VS. HOST DISEASE IN MICE

VAN RAFFARD-VAN DER VEEN F M; KONG Y M; ROSE N R; KIMURA M; GLEICHMANN E
PUBLICATION SECRETARIAT, CENTRAL LAB. NETHERLANDS RED CROSS BLOOD
TRANSFUSION SERV., P.O. BOX 9506, 1006 AK AMSTERDAM, NETHERLANDS.

CLIN EXP IMMUNOL 55 (3). 1984. 525-534. CODEN: CEXIA

Full Journal Title: Clinical and Experimental Immunology

Language: ENGLISH

In contrast to lupus-like autoantibodies such as anti-DNA, autoantibodies to mouse thyroglobulin (MTg) were not previously detectable in serum of F1 mice suffering from a lupus-like graft vs. host disease (GVHD) (GVH F1). Possible explanations for this restricted autoantibody formation during the potent allogeneic stimulation were investigated. The main question was whether the natural level of circulating MTg was too low to induce the formation of anti-MTg antibodies in GVH F1 mice. Existence, in the F1 mice studied, of B cells capable of producing anti-MTg antibodies was demonstrated by injection of lipopolysaccharide (LPS) and exogenous MTg. MTg injected into various F1 mice at the onset of the GVH reaction (GVHR) failed to overcome the lack of antibody formation to MTg even though the GVHR led to a severe lupus-like disease. Adult thymectomy (ATx) of either the recipients, the donors or both also did not break tolerance to MTg during the GVHR, irrespective of administration of exogenous MTg. Thus, neither i.v. injection of MTg nor ATx, designed to remove T suppressor (Ts) cells, is adequate to enable an autoantibody response to MTg during lupus-like GVHD. Hence, the non-specific T cell help that causes lupus-like GVHD seems to be intrinsically insufficient to trigger the Tg reactive B cells. Globular proteins, such as Tg, may require specific T cell help. In the presence of only non-specific T help, self-antigens such as DNA seem to be more apt than globular proteins to provide an effective signal 1 to the corresponding autoreactive B cells.

7/7/5 (Item 5 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

4408376 BIOSIS Number: 77083703

NATURE OF HISTAMINE INDUCED SUPPRESSOR CELLS ON THE EFFECTOR PHASE OF DELAYED HYPER SENSITIVITY TO DI NITROFLUORO BENZENE IN MICE

SATO K

THIRD DEP. INTERNAL MED., SCH. MED., UNIV. TOKUSHIMA, TOKUSHIMA.

SHIKOKU ACTA MED 39 (3). 1983 (RECD. 1984). 291-300. CODEN: SKIZA

Full Journal Title: Shikoku Acta Medica

Language: JAPANESE

The nature and role of suppressor cells on the effector phase of delayed

type hypersensitivity (DTH) to 2,4-dinitro-1-fluorobenzene (DNFB) were investigated in mice using cell transfer experiment. BALB/c mice were sensitized by high dose (2.5%) DNFB, and spleen cell suspension was made 5 days later. These sensitized donor spleen cells were incubated with 10⁻⁴M histamine in vitro for 30 min at 37.degree. C, and were transferred i.v. to the 1% DNFB, and spleen cell suspension was made 5 days later. These sensitized donor spleen cells were incubated with 10⁻⁴ M histamine in vitro for 30 min at 37.degree. C, and were transferred i.v. to the 1% DNFB sensitized recipient mice. DTH was measured by ear swelling 1 day after challenge of 1% DNFB. Suppression induced by the treatment with histamine was specific to the antigen DNFB. T-suppressor cells (Ts-amp cells) in the spleen were sensitive to anti-Thy-1 antibody. These cells could also be passed through a nylon wool column. Suppression disappeared by pretreatment of donor mice with cyclophosphamide (CY). Thus, histamine induced DTH Ts-amp cells in the spleen are sensitive to CY. The spleen cells from donor mice thymectomized at adult (ATX) before sensitization with DNFB did not express suppressive activity on DTH. The histamine-receptor positive Ts-amp cells were present in the spleen cells of mice which were sensitized to a large dose of DNFB and induced suppressor T cells to regulate the effector phase of DTH in the presence of histamine. The role of the suppressor cells was discussed.

7/7/6 (Item 6 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

4084272 BIOSIS Number: 76034123

CHARACTERIZATION OF A CONCAVALIN A INDUCED AMPLIFIER T CELL WHICH AUGMENTS IN-VITRO ANTIBODY RESPONSES TO DI NITRO PHENYL FICOLL

FINESILVER A G; BRALEY-MULLEN H

DEP. MICROBIOL., UNIV. MO., COLUMBIA, MO. 65212.

CELL IMMUNOL 75 (2). 1983. 199-213. CODEN: CLIMB

Full Journal Title: Cellular Immunology

Language: ENGLISH

The addition of the T-cell mitogen concanavalin A (Con A) on Day 2 of a 4-day in vitro culture of murine spleen cells with the thymus-independent (TI) antigen DNP[dinitrophenyl]39-Ficoll resulted in significant enhancement of the direct antitrinitrophenyl (TNP) plaque-forming cell (PFC) response. This enhancement was mediated by a nylon wool- and antiIg-nonadherent amplifier T cell (TA). TA activity was not eliminated by in vitro treatment of T cells with anti-Thy 1.2 and complement. TA activity could be eliminated by pretreatment of mice with antilymphocyte serum (ALS) in vivo, followed in vitro treatment of T cells with anti-Thy 1.2 + C. TA appear to bear a low surface density of Thy-1 antigen. These TA were relatively resistant to ALS used alone, to cyclophosphamide, and to low dose in vitro irradiation. TA were still present in the spleen 14 wk after adult thymectomy (ATx). They were I-J positive and apparently belonged to the Lyt 1+2- T-cell subset.

7/7/7 (Item 7 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

3986737 BIOSIS Number: 75034096

CORTICO STEROID ACTION ON LYMPHOCYTE SUB POPULATIONS AND HUMORAL IMMUNE RESPONSE OF AXOLOTL URODELE AMPHIBIAN

TOURNEFIER A

LAB. D'IMMUNOL. COMPAREE, UNIV. PIERRE MARIE CURIE, 9 QUAI SAINT-BERNARD, 75005 PARIS, FR.

IMMUNOLOGY 46 (1). 1982. 155-162. CODEN: IMMUA

Full Journal Title: Immunology

Language: ENGLISH

The effect of in vivo hydrocortisone (HC) treatment on thymocytes, splenic and blood lymphocytes and on allogeneic and humoral immune

responses were investigated in the axolotl [*Ambystoma mexicanum*]. HC induces a profound lymphocytopenia in the thymus (83% HC sensitive) and the spleen (50% HC sensitive) but not in the blood. The density gradient analysis of HC-treated axolotls showed that thymic cell populations of light density were more sensitive than populations of high density. The timing of HC administration in relation to the antigenic challenge is crucial for the humoral immune response. If HC injection is given 8 days before or on the same day as injection of horse red blood cells (HRBC), the antibody response is markedly enhanced. If HC injection occurs 8 days after injection of HRBC or on the day of the maximum anti-HRBC response, the antibody response is unchanged. The allograft immune response is not affected by HC treatment. The parallelism of the enhanced anti-HRBC response after HC treatment, with the same enhanced response obtained in adult thymectomized (ATx) animals, as well as in ATx-HC-treated axolotl, may be explained by the presence of a corticosterone-sensitive, T-like population of suppressor cells in axolotl.

7/7/8 (Item 8 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

3661783 BIOSIS Number: 73054150

SERUM THYMIC FACTOR RESTORES THE ABILITY OF ADULT THYMECTOMIZED MICE TO BE SUPPRESSED BY HAPTEN MODIFIED SELF

GROUX B; CHARREIRE J; ERARD D; GALANAUD P; BACH J-F
INSERM U131, 32 RUE DES CARNETS, 92140 CLAMART, FR.
CELL IMMUNOL 64 (1). 1981. 144-149. CODEN: CLIMB
Full Journal Title: Cellular Immunology
Language: ENGLISH

The injection of trinitrophenyl (TNP) conjugated spleen cells (TNP-SC) into normal recipients induces a specific suppression of the anti-TNP antibody response to the T-independent antigen TNP polyacrylamide. The T-cell dependency of this suppression can be shown by 2 arguments. Splenic T cells from TNP-SC-treated mice may suppress the in vitro antibody response of normal mouse spleen cells. The anti-TNP response of adult thymectomized (ATX) mice is evidently not suppressed after injection of TNP-SC. Treatment of such ATX mice with a synthetic analog of serum thymic factor (FTS) 7-9 wk after thymectomy restores their susceptibility to TNP-SC-induced suppression.

7/7/9 (Item 9 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

3398958 BIOSIS Number: 72031349

POLY GENIC CONTROL OF THE IMMUNE RESPONSE TO F ANTIGEN
SILVER D M; LANE D P

SLOAN KETTERING CANCER CENT., 1250 FIRST AVE., BOX 41, NEW YORK, N.Y.
10021.

IMMUNOGENETICS 12 (3-4). 1981. 237-252. CODEN: IMNGB
Full Journal Title: Immunogenetics
Language: ENGLISH

The ability to produce an autoimmune response to F antigen in mice is under H-2-linked and non-H-2-linked Ir-gene control. There is an absolute requirement for a k allele at H-2K or I-A to produce anti-F antibodies. Low and high responsiveness is controlled by a non-H-2-linked Ir gene which behaves in a similar fashion to Ir-3, in that as the dose of F-antigen is lowered, low responders behave as high responders and vice versa. This conversion from low to high responsiveness also occurs within a month after ATX [adult thymectomy]. Most F1 hybrids derived from (responder . times. nonresponder) parents bearing identical F-types behave as dominant nonresponders. As a result of ATX, such F1 mice convert to high responders. This conversion occurs if the animals are not immunized before day 90. If they receive F antigen prior to that time, they remain nonresponders for

7-9 mo. One F1 combination, AKD2, behaves as a dominant high responder. Genetic analysis showed that in the presence of a k allele at H-2K or I-A, a non-H-2-linked Ir gene inherited from the AKR mice determined dominant responsiveness. No manipulation of the immune response or combination of genes converted nonresponders lacking a k allele into responders. Such complex genetic control suggests regulation by a number of independently segregating loci whose function it is to limit the autoimmune response to F antigen.

7/7/10 (Item 10 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

3128622 BIOSIS Number: 70078529

THYMUS DEPENDENT SUPPRESSION OF HELPER FUNCTION IN ADULT XENOPUS-LAEVIS
THE SOUTH AFRICAN CLAWED TOAD

RUBEN L N; METTE S A; COCHRAN S K; EDWARDS B F
DEP. BIOL., REED COLL., PORTLAND, OREG. 97202, USA.
THYMUS 2 (1). 1980. 19-26. CODEN: THYMD
Full Journal Title: Thymus
Language: ENGLISH

Carrier-dependent and specific helper function capable of amplifying anti-hapten (TNP [trinitrophenyl]) responses can be demonstrated in *X. laevis*. This amplifying activity appears to be short-lived, since separation of the carrier (RBC [red blood cell]) priming and hapten-carrier injections by only 1 wk eliminates the effect of carrier priming. Carrier enhancement is optimal when the injections are made 2 to 4 days apart. Surgical removal of both thymuses (ATx) of young adults 2 days after carrier priming allows helper function to persist for at least 2 wk. The magnitude of the anti-hapten response in ATx animals is substantially greater than in sham thymectomized (STx) siblings, even when an optimal immunization schedule was utilized. Thymus-dependent suppression of an anti-heterologous erythrocyte response was demonstrated in vitro. Animals were immunized in vivo and their individual thymuses and spleen halves were subsequently cultured separately and together. The addition of thymus with its corresponding spleen substantially reduced the number of splenic antigen-binding cells and the titer of antibody released into the culture medium. This relatively primitive amphibian apparently possesses thymus-derived suppressor cells which regulate humoral immune responses.

7/7/11 (Item 11 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

3027463 BIOSIS Number: 69064870

ACCELERATION OF AGE RELATED CHANGES IN THE IMMUNITY AND ENDOCRINE ORGANS
BY ADULT THYMECTOMY IMMUNOLOGICAL AND HISTO PATHOLOGICAL STUDIES

HAYASHI Y; HIROKAWA K
DEP. PATHOL., MED. RES. INST., TOKYO MED. DENT. UNIV., 5-45 YUSHIMA,
BUNKYO, TOKYO 113, JPN.
ACTA PATHOL JPN 29 (6). 1979 (RECD. 1980). 933-948. CODEN: APJAA
Full Journal Title: Acta Pathologica Japonica
Language: ENGLISH

Long-lived B6C3F1 female mice were adult thymectomized (aTX) at 6 wk of age, and the effect of aTX was observed from immunological and histopathological viewpoints up to 24 mo. of age. Anti-SRBC [sheep red blood cells] antibody response in aTX group showed an apparently accelerated decline with age as compared with age-matched control. Mitogenic responsiveness of spleen and lymph node cells to phytohemagglutinin (PHA) and concanavalin (Con) A in the aTX group showed a slightly accelerated decline with age, although statistically it was not significant except for the responsiveness of spleen cells to PHA [phytohemagglutinin]. Thymic function to induce helper T [thymus-derived] cells was maintained throughout life, although it declined with age, but

its function to induce PHA- and Con A-responsive T cells disappeared in an early stage of life. Histopathologically, age related changes of the endocrine organs were apparently accelerated in the aTX group, suggesting that the normal function of endocrine organs appeared to have correlation with normal function of the thymus throughout life. The aTX increased the incidence of spontaneously occurring reticulum cell sarcoma in aged mice, supporting the concept of immunological surveillance.

7/7/12 (Item 12 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

2468000 BIOSIS Number: 66014905

THE SUPPRESSION OF THE DEVELOPMENT OF EXPERIMENTAL THYROIDITIS AND
RELATED CELLULAR IMMUNITY IN THE GUINEA-PIG FOLLOWING ADULT THYMECTOMY
HOJO K; HIRAMINE C

DEP. PATHOL., OSAKA CITY UNIV. MED. SCH., OSAKA 545, JPN.

TOHOKU J EXP MED 124 (3). 1978 251-260. CODEN: TJEMA

Full Journal Title: Tohoku Journal of Experimental Medicine

Language: ENGLISH

The role of the thymus in the pathogenesis of experimental allergic thyroiditis (EAT) in the guinea pig was studied after the lapse of a sufficiently long period following adult thymectomy. Female Hartley guinea pigs thymectomized or sham-operated at 10-12 wk of age were immunized by a single injection of homologous thyroid extract in complete Freund's adjuvant (CFA) 7 or more mo. after operation. The animals were sacrificed at 4 or 5 wk after immunization. Adult thymectomized and subsequently sensitized (ATx-sensitized) animals showed a markedly depressed ability to develop thyroiditis, 18 of the 21 animals thyroid lesions which could be graded as slight or lower than sham-operated controls. Delayed skin reaction and macrophage migration inhibitory factor production of lymph node cells to thyroid antigen were reduced in the ATx-sensitized animals. Enhancement of migration was noted in some of the animals. Anti-thyroid hemagglutinating antibody was not detected in about a quarter of ATx-sensitized animals, in the remainders of which the titers were at the level similar to that of the sham-operated controls. There was a substantial decrease in the percentage of rosette-forming cells (T [thymus derived] cells) on unsensitized ATx animals 14 mo. after operation. The suppression of the development of thyroiditis and related cellular immunity may be a reflection of a decline in peripheral T cell population during the long term after thymectomy. EAT in the guinea pig may be a thymus dependent disease.

7/7/13 (Item 13 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

2444442 BIOSIS Number: 65070850

SUB POPULATIONS OF SPLENIC THYMUS DERIVED CELLS REGULATING AN ANTI HAPTEN
ANTIBODY RESPONSE PART 1 HELPER AND AMPLIFIER CELLS

MUIRHEAD D Y; CUDKOWICZ G

DEP. PATHOL., SCH. MED., STATE UNIV. N.Y., BUFFALO, N.Y. 14214, USA.

J IMMUNOL 120 (2). 1978 579-585. CODEN: JOIMA

Full Journal Title: Journal of Immunology

Language: ENGLISH

There was a pronounced quantitative difference between the helper activities of B6C3F1 splenic T [thymus derived] cells sensitized with unmodified vs. modified antigens of SRBC [sheep red blood cells]. Modified SRBC induced the greater helper activity which was measured by the magnitude of an anti-TNP [trinitrophenol] response (Ig[immunoglobulin]M and IgG) elicited in vivo by virgin B [bone marrow derived] lymphocytes. Antigen modification was produced by conjugating SRBC with the hapten or simply by incubating SRBC in cacodylate buffer. There were restrictions with respect to erythrocyte species and mouse strains for this differential

priming to occur. The relatively poor performance of SRBC-primed T lymphocytes was apparently not due to suppressor T cells, but rather to activation of only 1 of 2 identified T cell subpopulations required for full helper activity. Unmodified SRBC activated a subpopulation of helper cells characterized as sensitive to elimination by ATS [anti-thymocyte serum] and long-lived after ATx [adult thymectomy] but failed to activate in B6C3F1 mice a second subpopulation of amplifier cells resistant to elimination by ATS and short-lived after ATx. Modified SRBC activated helper and amplifier cells. Under appropriate conditions these subsets of T cells were strongly synergistic in promoting anti-hapten antibody formation especially of the IgG class. The involvement of 2 distinct types of T lymphocytes in the positive regulation of antibody responses raises interesting and novel questions concerning the sequence of events in the triggering of B cells and the subsequent development of the response.

7/7/14 (Item 14 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

2426504 BIOSIS Number: 65052912

REGULATION OF ANTI HAPTEN ANTIBODY RESPONSE BY CHEMICALLY MODIFIED CARRIER ANTIGEN PREFERENTIALLY PROVOKING DELAYED TYPE HYPER SENSITIVITY PART 2 DELAYED TYPE HYPER SENSITIVITY REACTIVITY AND CARRIER SPECIFIC SUPPRESSION OF ANTI DI NITRO PHENOL ANTIBODY RESPONSE INDUCED BY PRIMING WITH DODECANOYL BOVINE SERUM ALBUMIN ARE MEDIATED BY FUNCTIONALLY DISTINCT THYMUS DERIVED CELL SUB POPULATIONS

MACHIDA A; KUMAZAWA Y; MIZUNOE K

DEP. IMMUNOL., KITASATO INST., 5-9-1 SHIROKANE, MINATO, TOKYO 108, JPN.

MICROBIOL IMMUNOL 21 (8). 1977 439-450. CODEN: MIIMD

Full Journal Title: Microbiology and Immunology

Language: ENGLISH

Experiments were carried out to determine whether the cell populations involved in DTH and the suppression of antibody response are identical. The effects of 4 treatments, i.e., adult thymectomy (ATx), X-irradiation, anti-mouse thymocyte serum (ATS) and hydrocortisone (HC) on the induction of DTH and on the carrier-specific suppression of antibody response were observed in mice immunized with chemically modified antigen, dodecanoyl-BSA [bovine serum albumin] (d-BSA), emulsified with complete Freund's adjuvant (CFA). DTH induced by immunization with d-BSA remained constant in adult thymectomized mice, whereas the suppression of antibody response was not inducible in these animals. Injection of low doses of ATS caused the depression of DTH in mice primed with d-BSA, but did not affect the suppressive activities of their spleen cells. Sublethal X-irradiation 1 wk prior to d-BSA priming inhibited the generation of suppressor cells but did not affect the generation of cells mediating DTH. The suppressive effect was also abrogated by sublethal X-irradiation given 2 days after immunization with DNP[dinitrophenylated]-BSA (14 days after priming with d-BSA). The treatment of animals with HC 2 days before the footpad challenge or immunization with DNP-BSA depressed the ability of animals to induce DTH and the suppression of antibody response. However, the latter was more sensitive to HC than the former. d-BSA-primed spleen cells were capable of suppressing anti-DNP response, but not of inducing DTH-reactivity upon transfer to recipient mice. DTH-reactivity and the carrier-specific suppression of anti-hapten antibody response induced by injection of d-BSA are apparently mediated by different cell populations.

7/7/15 (Item 1 from file: 434)
DIALOG(R)File 434:Scisearch(R)
(c) 1994 Inst for Sci Info. All rts. reserv.

12637293 Genuine Article#: MB824 Number of References: 26

Title: ENHANCED MYOCARDIAL LESIONS IN CHRONICALLY

TRYPANOSOMA-CRUZI-INFECTED RATS SUBJECTED TO ADULT THYMECTOMY

Author(s): BOTTASSO OA; REVELLI SS; DAVILA H; VALENTI JL; MUSSO OC; FERRO

ME; ROMEROPIFFIGUER M; MORINI JC

Corporate Source: UNIV NACL ROSARIO, FAC CIENCIAS MED, DIV INMUNOL, SANTA FE
3100/RA-2000 ROSARIO//ARGENTINA/; UNIV NACL ROSARIO, FAC CIENCIAS
MED, DEPT PATOL/RA-2000 ROSARIO//ARGENTINA/; NATL UNIV CORDOBA, FAC
CIENCIAS QUIM/RA-5000 CORDOBA//ARGENTINA/

Journal: IMMUNOLOGY LETTERS, 1993, V37, N2-3 (AUG), P175-180
ISSN: 0165-2478

Language: ENGLISH Document Type: ARTICLE

Abstract: Control animals and rats infected 90 days earlier, by inoculation of 1×10^6 trypomastigotes of *Trypanosoma cruzi* at weaning, were subjected to adult thymectomy (ATx) or sham operation (S-ATx) and assessed 3 months later for the presence of myocardial lesions and levels of lymph node and spleen T-cell populations. Chronic focal myocarditis (CFM) developed in 78% and 84% of S-ATx or ATx infected rats, respectively. While the two groups of infected rats did not differ as to the occurrence of myocardial lesions, large foci of CFM were more prevalent in ATx infected rats. Chronic *T. cruzi* (Tc) infection resulted in decreased CD4+ and increased CD8+ lymph node and spleen cell, with CD8+ lymphocytes being lowered to normal values in the spleen of the ATx infected group. It is suggested that ATx might act by interfering with a down-regulating immunoregulatory mechanism, leading to an exacerbation of autoimmune reactions believed to be involved in the generation of myocardial damage.

7/7/16 (Item 1 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

07076118 89378118

Islet cell transplantation in type I diabetes mellitus: evaluation of humoral immune response.

Kondratiev YY; Sadovnikova NV; Petrova GN; Fedotov VP; Bljumkin VN; Ignatenko SN; Pankov YA

Institute for Experimental Endocrinology and Hormone Chemistry, Academy of Medical Sciences, Moscow, USSR.

Exp Clin Endocrinol (GERMANY, EAST) May 1989, 93 (2-3) p147-50, ISSN 0232-7384 Journal Code: EPA

Languages: ENGLISH

Document type: JOURNAL ARTICLE

Four males and three females ranging in age from 20 to 35 years and afflicted with complicated Type 1-diabetes for more than 8 years underwent islet cell allotransplantation (ATx, 6 cases) and xenotransplantation (XTx, 1 case). Precultured islet cells derived from human or bovine fetal pancreata were injected into the m. rectus abdominis. Immunosuppression was not applied. Plasma C-peptide and islet cell surface antibodies (ICSA) were continually measured both before and until the twentieth week following islet cell transplantation. All recipients were subdivided as "responsive" (RR, 3 males) or "non-responsive" (NRR, 1 male and 3 females), according to the dynamics of their ICSA levels. All 3 RR (1XTx and 2 ATx) showed a peak of ICSA two weeks after cell injection. Subsequent ICSA levels had the tendency to either diminish or increase. Heterogeneity of preoperative antibody level, especially in NRR, was also observed. No associations between ICSA and ATx or XTx, age at diabetes onset, or duration of the disease was found. Only one RR with XTx had a reduced daily insulin requirement and a significant C-peptide response similar to the dynamics of ICSA levels. A greater mass of available bovine islet cells might be responsible for this effect.

7/7/17 (Item 2 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

05869676 86170676

Temporal changes of suppressor T lymphocytes and cytotoxic T lymphocytes

in syngeneic murine malignant gliomas.

Yamasaki T; Handa H; Yamashita J; Namba Y; Hanaoka M

J Neurooncol (UNITED STATES) 1986, 3 (4) p353-62, ISSN 0167-594X

Journal Code: JCF

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The temporal activities of suppressor T lymphocytes (Ts) and cytotoxic T lymphocytes (CTL) were investigated in a syngeneic murine malignant glioma (a methylcholanthrene-induced ependyoblastoma of C57BL/6 mouse origin, 203-glioma). After the s.c. tumor inoculation, it was suggested that both Ts and CTL were generated with target specificity against 203-glioma cells, because neither Ts nor CTL activity were seen against syngeneic EL 4 (benzpyrene-induced thymoma), allogeneic P815 (methylcholanthrene-induced mastocytoma of DBA/2 mouse origin) or YAC-1 (Moloney leukemia-induced T-cell lymphoma of A/Sn mouse origin), but only against 203-glioma. It was found that the generation of Ts preceded that of CTL and that the turnover was faster; furthermore, Ts were generated in the thymus and spleen, while CTL were distributed in regional lymph nodes and spleen. Surface marker analysis revealed that only Lyt-1-.2.3+ T-cells participated in suppressor responses in contrast to both Lyt-1-.2.3+ and Lyt-1+.2.3+ T-cells participating in cytotoxic responses. The effects of adult thymectomy (ATx) on the changes of the immunized T-cell subsets were also investigated. In mice thymectomized 3 weeks previously, the Ts activity was abrogated, whereas the CTL activity increased markedly and Lyt-1+.2.3+ T-cells were not detected. The results suggest that CTL or their precursors bearing Lyt-1+.2.3+ phenotype and Ts bearing Lyt-1-.2.3+ phenotype are short-lived lymphocytes. Accordingly, it is suggested that in tumor-bearing mice short-lived Ts are generated earliest with target specificity and, due to the reciprocal relationships between Ts and CTL activities, may have a modulating influence on CTL; furthermore, ATx may alter the patterns of generation of the precursor T-cells and Ts.

7/7/18 (Item 3 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

05389460 85005460

T-cell recruitment regulated by prostaglandin-mediated system and its role in immune response.

Koga Y; Taniguchi K; Nomoto K

Immunobiology (GERMANY, WEST) May 1984, 166 (4-5) p382-96, ISSN 0171-2985 Journal Code: GH3

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The dynamics of the number of T cells in spleen and the level of prostaglandin E2 in plasma were investigated serially in mice injected with Corynebacterium parvum. In the first few days, the level of plasma PGE2 increased but decreased to lower than the normal level thereafter. The absolute number of T cells in the spleen began to increase after the PGE2 level dropped. But such an increase of T cells was not observed in ATx mice challenged with C. parvum. Moreover, replenishing the mice with exogenous PGE2 in the period of low PGE2 halted selectively the increase of T cells in the spleen. This enlarged T cell subset responded to PHA, expressed Lyt-1+2+, and was sensitive to PGE2. And this T-cell subpopulation exerted a suppressive effect on antibody response in low PG environment, but lost its inhibitory effect in high PG milieu. These results suggested that an immature T cell subset is recruited from the thymus in a low PG state and participates as regulator cells in immune response at peripheral lymphoid organs.

7/7/19 (Item 4 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

05233266 84157266

Injection of mouse thyroglobulin and/or adult thymectomy do not break tolerance to thyroglobulin during the lupus like graft versus host disease in mice.

van Rappard-van der Veen FM; Kong YM; Rose NR; Kimura M; Gleichmann E
Clin Exp Immunol (ENGLAND) Mar 1984; 55 (3) p525-34, ISSN 0009-9104
Journal Code: DD7

Contract/Grant No.: AM 30975

Languages: ENGLISH

Document type: JOURNAL ARTICLE

In a previous paper (Gleichmann, van Elven & van der Veen, 1982), it had been reported that, in contrast to lupus like autoantibodies such as anti-DNA, autoantibodies to mouse thyroglobulin (MTg) were not detectable in serum of F1 mice suffering from a lupus like graft versus host disease (GVHD) (GVH F1). In the present paper, possible explanations for this restricted autoantibody formation during the potent allogeneic stimulation were investigated. The main question was whether the natural level of circulating MTg was too low to induce the formation of anti-MTg antibodies in GVH F1 mice. Existence, in the F1 mice studied, of B cells capable of producing anti-MTg antibodies was demonstrated by injection of lipopolysaccharide (LPS) and exogenous MTg. However, MTg injected into various F1 mice at the onset of the GVH reaction (GVHR) failed to overcome the lack of antibody formation to MTg even though the GVHR led to a severe lupus like disease. Furthermore, adult thymectomy (ATx) of either the recipients, the donors, or both also did not break tolerance to MTg during the GVHR, irrespective of administration of exogenous MTg. Thus, neither intravenous injection of MTg nor ATx, designed to remove T suppressor (TS) cells, is adequate to enable an autoantibody response to MTg during lupus like GVHD. Hence, the non-specific T cell help that causes lupus like GVHD seems to be intrinsically insufficient to trigger the Tg reactive B cells. We suggest that globular proteins, such as Tg, require specific T cell help. In the presence of only non-specific T help, self-antigens such as DNA seem to be more apt than globular proteins to provide an effective signal 1 to the corresponding autoreactive B cells.

7/7/20 (Item 5 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

05112117 84036117

Differentiation and maturation of thy-1 negative bone marrow cells. I. Effects of the thymus and radioresistant helper functions on the maturation of precursor cells specific for heterologous erythrocytes.

Gondo H; Taniguchi K; Kubo C; Nomoto K
J Clin Lab Immunol (ITALY) Sep 1983; 12 (1) p41-5, ISSN 0141-2760
Journal Code: J3K

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The effects of adult thymectomy (ATx) and preimmunization on the differentiation of cells responsible for delayed footpad reaction (DFR), plaque forming cells (PFC) and cell mediated lympholysis (CML) were examined in lethally irradiated and thy-1 negative bone marrow cell reconstituted C57BL/6 recipients. ATx reduced the degrees of immune responses in irradiated and reconstituted mice. When the recipients had been preimmunized, all of DFR, PFC and CML became detectable even in irradiated and reconstituted ATx mice. Preimmunization also evoked early maturation of precursor cells for CML. Therefore, it was suggested that radioresistant helper effects, presumably in the presence of antigens, could promote the differentiation and maturation of T cell precursors in bone marrow in the absence of the thymus. We also demonstrated differences in restoration periods of such responses after lethal irradiation and reconstitution. One or two weeks following irradiation and reconstitution, DFR was first detectable. On the other hand, the generation of PFC was detected later than 2 weeks after bone marrow cell reconstitution, and it took over 4 weeks for thy-1 negative bone marrow cells to raise CML. T

cells responsible for DFR may have lower dependency on the thymus than those for PFC and CML.

7/7/21 (Item 6 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

04922699 83155699

Characterization of a concanavalin A-induced amplifier T cell which augments in vitro antibody responses to DNP-Ficoll.

Finesilver AG; Braley-Mullen H
Cell Immunol (UNITED STATES) Feb 1 1983, 75 (2) p199-213, ISSN 0008-8749 Journal Code: CQ9
Contract/Grant No.: CA25054; AI-00322
Languages: ENGLISH
Document type: JOURNAL ARTICLE

The addition of the T-cell mitogen concanavalin A (Con A) on Day 2 of a 4-day in vitro culture of murine spleen cells with the thymus-independent (TI) antigen DNP39-Ficoll resulted in significant enhancement of the direct antitrinitrophenyl (TNP) plaque-forming cell (PFC) response. This enhancement was mediated by a nylon wool- and antiimmunoglobulin-nonadherent amplifier T cell (TA). TA activity was not eliminated by in vitro treatment of T cells with anti-Thy 1.2 and complement (C). TA activity could be eliminated by pretreatment of mice with antilymphocyte serum (ALS) in vivo, followed by in vitro treatment of T cells with anti-Thy 1.2 + C. Thus, TA appear to bear a low surface density of Thy-1 antigen. These TA were relatively resistant to ALS used alone, to cyclophosphamide, and to low dose in vitro irradiation. TA were still present in the spleen 14 weeks after adult thymectomy (ATx). They were I-J positive and apparently belonged to the Lyt 1+2- T-cell subset.

7/7/22 (Item 7 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

04312719 81140719

Polygenic control of the immune response to F antigen.
Silver DM; Lane DP
Immunogenetics (GERMANY, WEST) 1981, 12 (3-4) p237-51, ISSN 0093-7711
Journal Code: GI4
Contract/Grant No.: AI-13984
Languages: ENGLISH
Document type: JOURNAL ARTICLE

The ability to produce an autoimmune response to F antigen in mice is under H-2-linked and non-H-2-linked Ir-gene control. There is an absolute requirement for a k allele at H-2K or I-A in order to produce antiF antibodies. Low and high responsiveness is controlled by a non-H-2-linked Ir gene which behaves in a similar fashion to Ir-3, in that as the dose of F-antigen is lowered, low responders behave as high responders and vice versa. This conversion from low to high responders and vice versa. This conversion from low to high responsiveness also occurs within a month after ATX.-Most F1 hybrids derived from (responder X nonresponder) parents bearing identical F-types behave as dominant nonresponders. As a result of ATX, such F1 mice convert to high responders. This conversion occurs if the animals are not immunized before day 90. If they receive F antigen prior to that time, they remain nonresponders for 7-9 months. One F1 combination showed--AKD2--behaves as an dominant higher responder. Genetic analysis showed that the presence of a K allele at H-2K or I-A, a non-H-2-linked Ir gene inherited from the AKR mice determined dominant responsiveness. No manipulation of the immune response or combination of genes converted nonresponders lacking a k allele into responders. Such complex genetic control suggests regulation by a number of independently segregating loci whose function it is to limit the autoimmune response to F antigen.

7/7/23 (Item 8 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

03896586 80007586

Regulation of contact sensitivity to DNFB in the mouse: effects of adult thymectomy and thymic factor.

Erard D; Charreire J; Auffredou MT; Galanaud P; Bach JF

J Immunol (UNITED STATES) Oct 1979, 123 (4) p1573-6, ISSN 0022-1767

Journal Code: IFB

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The contact sensitivity response to DNFB is decreased after adult thymectomy (ATX). This response decreases to 50% of the control response of normal age-matched mice as soon as 3 weeks after ATX and is not further depressed 9 to 16 weeks after ATX. These results suggest that two T cell subsets of different lifespan are involved in the anti-DNFB response. A circulating thymic factor (FTS) is able to restore the contact sensitivity response to DNFB when injected 3 to 9 weeks after ATX but not 16 weeks later. By contrast, FTS has a depressive effect on the contact sensitivity response to DNFB of normal mice through a cyclophosphamide-sensitive T cell subset. These results suggest that FTS regulates DNFB contact sensitivity by acting on a cyclophosphamide-sensitive T cell subset, still present 9 weeks after ATX but absent after 16 weeks. Thus although the T cell defect, causing a depression of the contact sensitivity reaction to DNFB is quantitatively similar 3 and 16 weeks after ATX, its nature is probably different.

7/7/24 (Item 9 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

03829031 79206031

Masugi nephritis in T cell-depleted state. Influence of adult thymectomy and anti-thymocyte serum.

Oite T

Acta Pathol Jpn (JAPAN) May 1979, 29 (3) p333-45, ISSN 0001-6632

Journal Code: 1NE

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The effects of surgical removal of the thymus or the administration of antiserum to thymus-derived lymphocytes on the development of Masugi nephritis were investigated in Wistar rats. While thymectomy at weaning (ATx) had no significant effect on the humoral antibody response to injected rabbit IgG, repeated injections of rabbit anti-rat thymocyte serum (ATS) suppressed it remarkably. Both treatments did not show morphological evidences that glomerular inflammation and injury were suppressed in the autologous phase of Masugi nephritis. Glomerular lesions in rats receiving the nephrotoxic serum (NTS) one month after ATx (ATx-1+NTS) appeared to be severer in hypercellularity, mitotic counts, the amount of deposited fibrin-related substances and crescent formation than those in other nephritic groups. Morphological study of ATS+NTS rats revealed that the glomerular changes were nearly equal to those of NTS-injected control rats, in spite of markedly suppressed humoral response to injected rabbit IgG and the absence of host IgG along the glomerular capillary walls.

7/7/25 (Item 10 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

03625232 79002232

LT-Lymphocyte regulation of contact sensitivity: effect of thymectomy in

adult mice]

Regulation de l'hypersensibilite de contact par les lymphocytes T : effet de la thymectomie a l'age adulte chez la Souris.

Erard D; Charreire J; Auffredou MT; Galanaud P

C R Acad Sci Hebd Seances Acad Sci D (FRANCE) May 29 1978, 286 (21) p1539-42, ISSN 0567-655X Journal Code: C9C

Languages: FRENCH Summary Languages: ENGLISH

Document type: JOURNAL ARTICLE English Abstract

6 weeks after adult thymectomy (ATx) in the Mouse, the contact sensitivity reaction to dinitrofluorobenzene (DNFB) is enhanced. 4 to 7 months after ATx, this reaction is deeply, but incompletely depressed, whereas the concomitant antibody response is not affected. These results suggest that both a suppressor, and an amplifier, T lymphocytes, the life span of which is different after ATx, are involved in the regulation of contact sensitivity. The effect of circulating thymic factor on this reaction suggests that this factor acts exclusively, at least in short treatments, on the suppressor function.

7/7/26 (Item 11 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

03454356 78088356

Subpopulations of splenic T cells regulating an antihapten antibody response. I. Helper and amplifier cells.

Muirhead DY; Cudkowicz G

J Immunol (UNITED STATES) Feb 1978, 120 (2) p579-85, ISSN 0022-1767 Journal Code: IFB

Languages: ENGLISH

Document type: JOURNAL ARTICLE

There was a pronounced quantitative difference between the helper activities of B6C3F1 splenic T cells sensitized with unmodified vs modified antigens of SRBC. Modified SRBC induced the greater helper activity which was measured by the magnitude of an anti-TNP response (IgM and IgG) elicited in vivo by virgin B lymphocytes. Antigen modification was produced by conjugating SRBC with the hapten or simply by incubating SRBC in cacodylate buffer. There were restrictions with respect to both erythrocyte species and mouse strains for this differential priming to occur. The relatively poor performance of SRBC-primed T lymphocytes was apparently not due to suppressor T cells, but rather to activation of only one of two identified T cell subpopulations required for full helper activity. Unmodified SRBC activated a subpopulation of "helper" cells characterized as sensitive to elimination by ATS and long-lived after ATx, but failed to activate in B6C3F1 mice a second subpopulation of "amplifier" cells resistant to elimination by ATS and short-lived after ATx. In contrast, modified SRBC activated both helper and amplifier cells. Under appropriate conditions these subsets of T cells were strongly synergistic in promoting antihapten antibody formation especially of the IgG class. The involvement of two distinct types of T lymphocytes in the positive regulation of antibody responses raises interesting and novel questions concerning the sequence of events in the triggering of B cells and the subsequent development of the response.

7/7/27 (Item 12 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

03240596 77142596

Immunity to sporozoite-induced malaria infection in mice. I. The effect of immunization of T and B cell-deficient mice.

Chen DH; Tigelaar RE; Weinbaum FI

J Immunol (UNITED STATES) Apr 1977, 118 (4) p1322-7, ISSN 0022-1767 Journal Code: IFB

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The cellular basis of immunity to sporozoites was investigated by examining the effect of immunization of T and B cell-deficient C57BL/6N X BALB/c AnN F1 (BLCF1) mice compared to immunocompetent controls. Immunization of T cell-deficient (ATX-BM-ATS) BLCF1 mice with x-irradiated sporozoites did not result in the generation of protective immunity. The same immunization protocols protected all immunocompetent controls. In contrast, B cell-deficient (micron-suppressed) BLCF1 mice were protected by immunization in the majority of cases. The absence of detectable serum circumsporozoite precipitins or sporozoite neutralizing activity in the micron-suppressed mice that resisted a sporozoite challenge suggests a minor role for these humoral factors in protection. These data demonstrate a preeminent role for T cells in the induction of protective immunity in BLCF 1 mice against a *P. berghei* sporozoite infection.

7/7/28 (Item 13 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

02732639 75139639

Studies on the thymic dependence of the immunoglobulin classes of the mouse (38570).

Bankhurst AD; Lambert PH; Miescher PA

Proc Soc Exp Biol Med (UNITED STATES) Feb 1975, 148 (2) p501-4, ISSN 0037-9727 Journal Code: PXZ

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The influence of the thymus on serum immunoglobulin (Ig) concentration was studied by a comparison of serum Ig levels in congenitally athymic (nu/nu) mice versus control littermate heterozygotes and adult thymectomized, irradiated, bone marrow reconstituted mice (ATx plus B) versus adult thymectomized, irradiated mice reconstituted with bone marrow and thymus (ATx plus BT). IN THE FORMER GROUP IgG1, IgA, and IgG2a were 8%, 17% and 31% of controls. IgM levels were increased (340%) compared to controls. When ATx plus B mice were compared to controls. When ATx plus B mice were compared with nonirradiated controls significant depressions were noted in serum IgG1 and IgM. The only significant decrease in serum Ig levels between ATx plus B and ATx plus BT was in IgG1. These results are discussed in terms of the effects of thymic influence, residual T lymphocyte population differences between the two groups, and the effect of irradiation.

7/7/29 (Item 14 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

02708552 75115552

Cellular cooperation during in vivo anti-hapten antibody responses. III. The helper cell activity of activated thymocytes, of spleen cells treated with anti-theta serum, and of spleen cells from anti-thymocyte serum-treated or adult thymectomized donors.

Janeway CA Jr.

J Immunol (UNITED STATES) Apr 1975, 114 (4) p1408-14, ISSN 0022-1767 Journal Code: IFB

Languages: ENGLISH

Document type: JOURNAL ARTICLE

An adoptive secondary anti-2,4-dinitrophenyl (DNP) antibody response involving T-B cell collaboration has been studied. In particular, attempts have been made to affect the unexpectedly steep log dose-response curve obtained when graded numbers of helper cells are transferred to irradiated recipients given a fixed number of B cells (premium effect). A variety of means were used to alter helper cell activity, and this activity was then measured quantitatively, as was the ability of the helper cells present after these treatments to give a premium effect. It was shown that

activated T cells are approximately twice as active as spleen cells in helper activity and give a comparable premium effect. Graded doses of anti-theta serum plus complement markedly reduce the helper activity of spleen cells without affecting the premium effect given by the residual cells. Treatment of primed cell donors with limited doses of heterologous anti-mouse thymocyte serum (ATS) before transfer does not affect B cell activity, but readily inactivates helper cells, again without affecting the premium effect given by the residual cells. Adult thymectomy (ATx) of helper cell donors before priming with carrier led initially to increased helper activity relative to age-matched control donors. This increase may reflect the loss of nonspecific suppressor T cells from spleens shortly after ATx. Late after ATx, there was about a 2-fold decrease in helper activity, probably reflecting a loss of helper cell precursors. At no time was there any change in the premium effect. In view of the failure of any of the techniques used to abolish the premium effect given by helper cells in this response, it seems likely that this premium effect is due to the cooperative interaction of two very similar types of mature T cell. Alternatively, the premium effect observed here may result from the interaction of two activities of a single type of T cell which is mediated by independent factors.

7/7/30 (Item 1 from file: 159)
DIALOG(R)File 159:CANCERLIT(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

00915284 92683707 ICDB/92683707
IMMUNOHISTOCHEMICAL DETECTION OF BREAST CANCER MICROMETASTASES IN
PRETRANSPLANT BONE MARROW (MEETING ABSTRACT)
Franklin WA; Johnston CS; Williams S; Hami L; Jones RB; Bast RC; Shpall
EJ

Dept. of Pathology, UCHSC, Denver, CO 80262
Proc Annu Meet Am Assoc Cancer Res; 33:A1199 1992 ISSN 0197-016X
Languages: ENGLISH
Document Type: MEETING ABSTRACT

Detection of small numbers of breast cancer cells is important in assessing the efficacy of purging of tumor from bone marrow prior to autologous transplantation (ATx). We have used a mixture of monoclonal antibodies, including 260F9, 520C3, and 317G5 (Baxter Corp), as well as BRE-3 (Dr. R Ceriani) in a sensitive alkaline phosphatase immunohistochemical assay to identify breast tumor cells in bone marrow. In a series of experiments in which CAMA breast tumor cells were added to normal marrow, tumor cells were consistently detected in bone marrows at a concentration of one tumor cell per 10(6) normal bone marrow cells. Tumor cells were also detected in buffy coats of marrows from patients undergoing ATx. Among 51 specimens examined, tumor cells were found by immunohistochemical staining of buffy coats in six specimens which had been tumor-negative by conventional light microscopy. Immunohistochemical staining of bone marrow buffy coats may more accurately reflect the presence or absence of tumor in bone marrow than conventional light microscopic examination of bone core sections.

7/7/31 (Item 2 from file: 159)
DIALOG(R)File 159:CANCERLIT(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

00186025 78706634 CARC/78706634
EFFECT OF ADULT THYMECTOMY ON TUMOUR IMMUNITY IN MICE
Takei F; Levy JG; Kilburn DG
Dept. Microbiology, Univ. British Columbia, Vancouver, British Columbia,
Canada V6T 1W5
Br J Cancer; 37(5):723-731 1978 ISSN 0007-0920
Languages: ENGLISH
Document Type: JOURNAL ARTICLE
The effect of adult thymectomy (ATx) on (1) the growth of the P815

mastocytoma in DBA/2J mice, (2) the ability of mice to generate cytotoxic cells specific for P815, and (3) the ability of tumor-bearing animals to generate suppressor cells specific for this tumor was studied. Spleen cells from adult mice that had been Tx 8wk previously demonstrated a severely impaired primary cytotoxic response to P815 tumor cells, whereas their cytotoxic responses to allogeneic cells (C57BL/6) and to non-H-2 antigens (BALB/c) and their ability to form a primary antibody response to sheep RBC were unimpaired. Suppressor T cells, specific for P815 cells, appeared after 4-8 days in the thymuses of animals inoculated with P815 cells. No differences in tumor growth between ATx animals and sham-operated controls were observed, and the Tx tumor-bearing animals and untreated controls had equal levels of specific suppressor cells in their lymph nodes. Severely thymocyte-deprived mice that had been Tx, irradiated, and reconstituted with either marrow or spleen cells 8 wk before tumor implantation succumbed more rapidly to metastatic tumor than did all control animals. The data suggest that tumor immunity and temporary tumor containment, at least with this tumor line, depend on the presence of an intact immune system. (20 Refs)

7/7/32 (Item 1 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119218353 CA: 119(21)218353q PATENT
Autotaxin: motility stimulating protein useful in cancer diagnosis and therapy.
INVENTOR(AUTHOR): Stracke, Mary; Liotta, Lance A.; Schiffmann, Elliott; Kratzsch, Henry
LOCATION: USA
ASSIGNEE: United States Dept. of Health and Human Services
PATENT: U.S. Pat. Appl. ; US 822043 A0 DATE: 930101
APPLICATION: US 822043 (920117)
PAGES: 61 pp. Avail. NTIS Order No. PAT-APPL-7-822 043. CODEN: XAXXAV
LANGUAGE: English
SECTION:
CA202010 Mammalian Hormones
CA201XXX Pharmacology
IDENTIFIERS: autotaxin human autocrine purifn
DESCRIPTORS:
Nomenclature, new natural products...
autotaxin (human cell motility-stimulating autocrine)
Lymphokines and Cytokines...
autotaxin, purifn. and characterization of human
Gene, animal...
cDNA, for human autotaxin, cloning of
Neoplasm...
diagnosis of, human autotaxin in relation to
Neoplasm inhibitors...
human autotaxin purifn. and characterization and autotaxin cDNA cloning in relation to
Amino acids, biological studies...
of human autotaxin
Molecular cloning...
of human autotaxin cDNA
Antibodies...
to human autotaxin
CAS REGISTRY NUMBERS:
147960-51-8 147960-52-9 147960-53-0 147960-54-1 147960-55-2
147960-56-3 147960-57-4 147960-58-5 147960-59-6 147960-60-9
147960-61-0 147960-62-1 147960-63-2 147977-73-9 150236-72-9
150979-01-4 150979-02-5 150979-03-6 150979-04-7 fragment of human autotaxin

7864340 EMBASE No: 90300650

Immune competence in 90Sr-exposed, adult thymectomized and antilymphocyteglobulin-treated CBA mice. II. Reticuloendothelial phagocytic function and in vitro mitogen responsiveness of spleen cells

Bierke P.

The Unit of Experimental Pathology, Faculty of Veterinary Medicine, Swedish University of Agricultural Sciences, Uppsala Sweden

ACTA ONCOL. (Sweden), 1990, 29/5 (615-621) CODEN: ACTOE ISSN: 0284-186X

LANGUAGES: English

The hypothesis that immune system failure plays a role in the development of radiation induced tumours was recently investigated experimentally. Young adult CBA mice, intact or immunocompromised by adult thymectomy (ATx) and/or antilymphocyteglobulin (ALG) treatment, were exposed to single doses of 90Sr, after which tumour development was monitored. To evaluate the experimental results required knowledge about the immunological responsiveness of the mice. The present paper contributes to that knowledge by reporting on the in vitro responsiveness of lymphoid cells to mitogens (LPS, PHA, Con-A) and the in vivo phagocytic functioning of the reticuloendothelial system (RES), measured as the rate of clearance of 125I-albumin micro-aggregates in peripheral blood. 90Sr, ATx and ALG-treatments, separately and in combination, suppressed mitogenic lymphoid cell activation, whereas the RES phagocytic function remained unchanged, except in response to 90Sr+ALG treatment, which seemed to slightly inhibit phagocytic activity.

7/7/34 (Item 2 from file: 73)

DIALOG(R)File 73:EMBASE

(c) 1994 Elsevier Science B.V. All rts. reserv.

7661047 EMBASE No: 90085328

Radiostrontium-induced oncogenesis and the role of immunosuppression. II. Influence of 90Sr dose, adult thymectomy and antilymphocyteglobulin treatment on the development of lympho-reticular and extraskkeletal neoplastic lesions in CBA mice

Bierke P.; Nilsson A.

Unit of Experimental Pathology and Risk Research, Department of Pathology, Faculty of Veterinary Medicine, Swedish University of Agricultural Sciences, S-750 07 Uppsala Sweden

ACTA ONCOL. (Sweden), 1990, 29/1 (53-63) CODEN: ACTOE ISSN: 0284-186X

LANGUAGES: English

The significance of depressed immune function for the development and progression of tumours induced by 90Sr (mainly osteosarcomas and malignant lymphomas) was investigated in a series of experiments by comparing the tumour responses in normal mice with those in immunocompromised mice. The present paper (part II) reports on lympho-reticular (LR) and extraskkeletal neoplastic lesions in male CBA/SU mice after exposure to different single doses of 90Sr with or without additional immunosuppression by adult thymectomy (ATx) and/or prolonged anti-lymphocyteglobulin (ALG) treatment. Neoplastic lesions in bone were reported in part I. The status of the animal's immune system and responsive ability were examined in parallel experiments. The tumour yields were analysed in relation to the dosage of 90Sr and the immunosuppressive treatments employed. Although the incidences and latency times of induced tumours were clearly dose-dependent, they were never significantly influenced by ATx/ALG treatments. Thus, no substantial support was gained for the theory that the immune system plays a controlling or modifying role in 90Sr carcinogenesis. The results, which are in agreement with the bone tumour responses, suggest that 90Sr induced tumours either do not express the antigens necessary for immune rejection or that the decline in immune responsiveness induced by ATx/ALG was of little consequence for tumour development and spread. The pathogenesis of 90Sr induced malignant lymphomas (MLs) and their immunophenotypes are

discussed.

7/7/35 (Item 3 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

7398664 EMBASE No: 89113667

Radiostrontium-induced oncogenesis and the role of immunosuppression. I. Influence of ^{90}Sr dose, adult thymectomy and antilymphocyteglobulin treatment on the development of neoplastic and preneoplastic lesions in the skeleton of CBA mice.

Bierke P.; Nilsson A.

Unit of Experimental Pathology and Risk Research, Department of Pathology, Faculty of Veterinary Medicine, Swedish University of Agricultural Sciences, S-750 07 Uppsala Sweden

ACTA ONCOL. (Sweden), 1989, 28/1 (87-102) CODEN: ACTOE ISSN: 0284-186X

LANGUAGES: English

Ionizing irradiation by incorporated strontium-90 exerts two major effects: it induces tumours (mainly osteosarcomas and lymphoreticular tumours) and depresses the immune system. The interrelation between these function, i.e. the significance of decreased immunological responsiveness in the oncogenic process, remains unclear. The influence of the ^{90}Sr dose and the role of immune modulation on the tumour yield, were investigated in young adult CBA mice. The animals were exposed to different single doses of ^{90}Sr and, in addition, some groups were subjected to long-term unspecific immune suppression by adult thymectomy (ATx) and/or prolonged antilymphocyteglobulin (ALG) treatment. The present paper (part I) reports on the effects of the treatments on bone tumour responses as reflected by incidence, multiplicity, latency time, histologic characteristics and growth behaviour. The histogenesis of osteosarcomas, as evidenced morphologically by preneoplastic and early neoplastic growth, is illustrated and discussed. The results demonstrate a positive dose-response relationship for osteosarcomas, in which the relative incidences of the various osteosarcoma subtypes were differentially affected. Thus, well-differentiated tumours were gradually replaced by less differentiated types as the dose decreased. A correlation was also observed between the incidence of osteosarcomas and that of assumed preneoplastic lesions in the same bones and sites. Immune suppression by ATx and/or ALG did not distinctly alter the neoplastic or preneoplastic responses at any dose-level of ^{90}Sr .

7/7/36 (Item 4 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

6190679 EMBASE No: 86185740

Influence of sup ^{90}Sr , adult thymectomy and antilymphocyteglobulin on haematopoietic tissues and peripheral blood leucocytes in CBA mice

Bierke P.

Unit of Radiologic Oncology, University of Stockholm, S-172 46 Sundbyberg SWEDEN

ACTA RADIOL. (SWEDEN), 1986, 25/2 (147-154) CODEN: AROBD

SERIES: SER. ONCOL.

LANGUAGES: ENGLISH

The role of long-time immune suppression in carcinogenesis induced by the long-lived emitter sup ^{90}Sr , is investigated in an ongoing study. The experimental design is based on the assumption that impaired immune responsiveness, by other means than sup ^{90}Sr , might increase the neoplastic response in exposed individuals, and thus reflect a protective function, if existing. Intercomparison is made of the tumour yield in mice exposed to different single doses of sup ^{90}Sr and simultaneously subjected or not to long-term immune suppression by adult thymectomy (ATx) and/or antilymphocyteglobulin (ALG) treatment. Information on the general

condition and responsiveness of the immune system, in the respective models, during tumour expectancy time, is essential for a conclusive evaluation of the results. To meet these demands that present paper reports on histopathologic alterations in immune organs and changes in white blood cell counts, induced by the different combinations of sup 9sup 0Sr, ATx and ALG treatment. The results confirm the prediction, that ATx + ALG is an efficient and, with respect to the purpose of the study, suitable treatment for additive long-term depression of the immune system in sup 9sup 0Sr irradiated mice, evidenced in particular by increased depletion of mononuclear cells (MNC) in lymphoid organs and peripheral blood. Subsequent reports will deal with functional immune parameters.

7/7/37 (Item 5 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

6016443 EMBASE No: 86011503

Cell-mediated lympholysis (CML) to allogeneic and trinitrophenyl (TNP)-modified cells: Re-evaluation of the role of the thymus in CML

Gondo H.; Tokuda N.; Taniguchi K.; Nomoto K.

Department of Immunology, Medical Institute of Bioregulation, Kyushu University 69, Fukuoka JAPAN

J. CLIN. LAB. IMMUNOL. (SCOTLAND), 1985, 17/3 (131-136) CODEN: JLIMD

LANGUAGES: ENGLISH

The role of the thymus in cell-mediated lympholysis (CML) to allogeneic and trinitrophenyl (TNP)-modified cells was re-evaluated using adult thymectomized (ATx) and neonatal thymectomized (NTx) mice. CML to TNP-modified cells was used as the model of the major histocompatibility complex (MHC)-restricted CML. CML to TNP-modified cells, but not to allogeneic cells was reduced at late stages after ATx and remained low even in NTx-11 (thymectomy at 11 days of age) mice. Interleukin-2 (IL-2) producing activity in such mice was lower than that in the normal controls, while allogeneic CML was maintained at the same level as seen in the controls. The addition of exogenous IL-2 to in vitro culture led to a restoration of the generation of CML to TNP-modified cells, in neonatal thymectomized mice. These results suggested that thymus-dependency differs in the 2 forms of CML and that the CML to TNP-modified cells showed a higher thymus-dependency than did the allogeneic CML. Moreover, the activity of helper T cells may exert a direct influence on CML to TNP-modified cells, as compared to allogeneic cells.

7/7/38 (Item 6 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

5631618 EMBASE No: 84127284

Influence of sup 9sup 0Sr, adult thymectomy and antilymphocyteglobuline on T-cells in mouse peripheral blood

Bierke P.; Gidlund M.

Unit of Radiologic Oncology, University of Stockholm, S-172 46 Sundbyberg SWEDEN

ACTA RADIOL. (SWEDEN), 1984, 23/1 (61-64) CODEN: AROND

SERIES: SER. ONCOL. RADIAT. THER. PHYS. BIOL.

LANGUAGES: ENGLISH

Three groups of male CBA mice were treated by (1) sup 9sup 0Sr injection (14.8 kBq/g body weight i.p.), (2) adult thymectomy (ATx) + antilymphocyteglobuline (ALG), and (3) sup 9sup 0Sr + ATx + ALG, respectively, and one untreated group served as a control. The relative number of T-lymphocytes was determined in peripheral blood mononuclear cells of individual animals using a dye exclusion cytotoxicity assay with antisera against the T-cell surface membrane marker Thy 1.2 and complement. Total and differential leucocyte counts were also performed. sup 9sup 0Sr irradiation decreased the total number of leucocytes irrespective of type, and the combined treatment of ATx and ALG decreased mainly mononuclear

cells and particularly T-cells. The most advanced T-cell depletion in peripheral blood was found in the sup 9sup 0Sr + ATx and ALG treated group with a 97 per cent reduction as compared with untreated controls. ATx + ALG thus proved to be useful for blood T-cell depletion in mice treated simultaneously with sup 9sup 0Sr, and might provide a valuable tool in investigations on the possible role of cell-mediated immune response in radiation-induced oncogenesis, with particular emphasis on selective depletion within the monomorphonuclear compartment.

7/7/39 (Item 7 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

2584389 EMBASE No: 81242583

Serum thymic factor (FTS) restores the ability of adult thymectomized mice to be suppressed by hapten-modified self

Grouix B.; Charreire J.; Erard D.; et al.

INSERUM U 13, 92140 Clamart FRANCE

CELL. IMMUNOL. (USA), 1981, 64/1 (144-149) CODEN: CLIMB

LANGUAGES: ENGLISH

The injection of trinitrophenyl (TNP) conjugated spleen cells (TNP-SC) into normal recipients induces a specific suppression of the anti-TNP antibody response to the T-independent antigen TNP polyacrylamide. The T-cell dependency of this suppression is shown by two arguments. (i) Splenic T cells from TNP-SC-treated mice suppress the in vitro antibody response of normal mouse spleen cells. (ii) The anti-TNP response of adult thymectomized (ATX) mice is not suppressed after injection of TNP-SC. Treatment of such ATX mice with a synthetic analog of serum thymic factor (FTS) 7-9 weeks after thymectomy restores their susceptibility to TNP-SC-induced suppression.

7/7/40 (Item 8 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

2064526 EMBASE No: 80201845

Trinitrobenzene sulfonic acid effects in two amphibian model systems

Mette S.A.; Ruben L.N.

Dept. Biol., Reed Coll., Portland, Ore. 97202 USA

CELL. IMMUNOL. (USA), 1980, 53/2 (298-306) CODEN: CLIMB

LANGUAGES: ENGLISH

The induction of hapten-specific tolerance was investigated in two amphibia, *Notophthalmus viridescens* and *Xenopus laevis*. Responses to trinitrophenylated (TNP)-Ficoll and TNP-lipopolysaccharide (LPS), as well as to horse erythrocytes (HRBC) were examined in both species, following an intraperitoneal injection of 2,4,6-trinitrobenzenesulfonic acid (TNBS). The less evolutionary advanced newt, *Notophthalmus*, failed to respond to all three immunogens after TNBS administration. While *Xenopus* became completely tolerant upon challenge with TNP-Ficoll and partially tolerant with TNP-LPS, full capacity to respond to HRBC was retained. Therefore, specific tolerance was induced in *Xenopus*, but not in *Notophthalmus*. The tolerance with TNP-Ficoll in the toad, *Xenopus*, was short lived and return to responsiveness appeared to be related inversely to levels of TNP protein in the sera of TNBS-treated animals. The thymic dependence of this tolerance could not be determined, because adult thymectomy (ATx) abrogated the response to TNP-Ficoll in control nontolerized animals. Responses to TNP-LPS and HRBC were unaffected by ATx. These data, in conjunction with TNBS-induced differential tolerance to the TNP moiety, suggest carrier-dependent hapten-specific B-cell heterogeneity in the toad which differs in certain ways from that recently described for murine systems.

7/7/41 (Item 9 from file: 73)
DIALOG(R)File 73:EMBASE

1051604 EMBASE No: 78223665

The involvement of Tsub 1 and Tsub 2 lymphocytes in primary and secondary delayed type hypersensitivity responsiveness

Van Der Kwast Th.H.; Benner R.

Dept. Cell Biol. Genet., Erasmus Univ., Rotterdam NETHERLANDS

ANN. IMMUNOL. (FRANCE), 1977, 128C/6 (1019-1024) CODEN: ANIMC

LANGUAGES: ENGLISH SUMMARY LANGUAGES: FRENCH

The contribution of short-lived and long-lived T lymphocytes (Tsub 1 and Tsub 2 lymphocytes respectively) to primary and secondary delayed type hypersensitivity (DTH) responsiveness was studied by means of thymectomy of adult mice (ATx) with sheep red blood cells as the antigen. Within 1 mth after ATx a fall of the primary and secondary DTH responsiveness was found of 20% and 50% respectively. This reduction can be attributed to the loss of short-lived T cells. ATx 8 mth prior to the administration of the priming dose caused a decrease of the primary and the secondary DTH responsiveness of about 50% and 65% respectively. These data suggest that both the short-lived Tsub 1 and the long-lived Tsub 2 precursor cells account for the occurrence of cells mediating the primary DTH response and memory cells responsible for the secondary DTH response. The contribution of both populations of precursor cells in primary and secondary DTH responsiveness appeared to be only proportionally different.

7/7/42 (Item 10 from file: 73)

DIALOG(R)File 73:EMBASE

(c) 1994 Elsevier Science B.V. All rts. reserv.

1051309 EMBASE No: 78223360

Regulation of anti-hapten antibody response by chemically modified carrier antigen preferentially provoking delayed-type hypersensitivity (DTH). II. DTH-reactivity and carrier-specific suppression of anti-DNP antibody response

Machida A.; Kumazawa Y.; Mizunoe K.

Dept. Immunol., Kitasato Inst., Tokyo JAPAN

MICROBIOL. IMMUNOL. (JAPAN), 1977, 21/8 (439-449) CODEN: MIIMD

LANGUAGES: ENGLISH

Experiments were carried out to determine whether or not the cell populations involved in DTH and in the suppression of antibody response are identical. The effects of 4 treatments, i.e., adult thymectomy (ATx), X irradiation, anti mouse thymocyte serum (ATS) and hydrocortisone (HC) on the induction of DTH and on the carrier specific suppression of antibody response were observed in mice immunized with chemically modified antigen, dodecanoyl BSA (d BSA), emulsified with complete Freund's adjuvant (CFA), with the following results: (1) DTH induced by immunization with d BSA remained constant in adult thymectomized mice, whereas the suppression of antibody response was not inducible in these animals. (2) Injection of low doses of ATS caused the depression of DTH in mice primed with d BSA, but did not affect the suppressive activities of their spleen cells. (3) Sublethal X irradiation 1 wk prior to d BSA priming inhibited the generation of suppressor cells but did not affect the generation of cells mediating DTH. The suppressive effect was also abrogated by sublethal X irradiation given 2 days after immunization with DNP BSA (14 days after priming with d BSA). (4) The treatment of animals with HC 2 days before the footpad challenge or immunization with DNP BSA depressed the ability of animals to induce both DTH and the suppression of antibody response. However, the latter was more sensitive to HC than the former. In addition to these results, it was also found that d BSA primed spleen cells were capable of suppressing anti DNP response, but not of inducing DTH reactivity upon transfer to recipient mice. These results suggest that DTH reactivity and the carrier specific suppression of anti hapten antibody response induced by injection of d BSA are mediated by different cell populations.

7/7/43 (Item 11 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

994026 EMBASE No: 78164251

Functional heterogeneity among the T derived lymphocytes of the mouse.
VII. Conversion of T1 cells to T2 cells by antigen

Araneo B.A.; Marrack P.; Kappler J.W.

Div. Immunol., Dept. Microbiol., Univ. Rochester Cancer Cent., Sch. Med.,
Univ. Rochester, N.Y. 14642 USA

J.IMMUNOL. (USA), 1977, 119/2 (765-771) CODEN: JOIMA

LANGUAGES: ENGLISH

The T1 subpopulation of peripheral T cells was defined in mice by its short half life, insensitivity to anti-thymocyte sera (ATS) in vivo, and slow kinetics of response to antigen. The T2 subpopulation was defined by its long lifetime, elimination by ATS in vivo, and rapid response to antigen. Mice containing only T1-type T cells were constructed by adult thymectomy (ATx) followed immediately by the elimination of T2 cells by ATS treatment. Immunization of these mice with SRBC led to the production of memory helper cells in the T2 subpopulation. This process depended on the presence of T1 cells and for the most part required SRBC immunization, although a few SRBC-specific T2 cells reappeared in the mice in the absence of antigen. We conclude that T1 cells can give rise to T2 cells in an antigen-driven step and that the 2 populations correspond to virgin and memory T cells, respectively.

7/7/44 (Item 12 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

884107 EMBASE No: 78050486

Biochemical characterisation of a serum thymic factor

Bach J.F.; Dardenne M.; Fleau J.M.; Rosa J.

INSERM U 25, Hop. Necker, Paris FRANCE

NATURE (LOND.) (ENGLAND), 1977, 266/5597 (55-57) CODEN: NATUA

LANGUAGES: ENGLISH

The authors presented direct evidence for the presence in normal blood of a thymus dependent nonapeptide of molecular weight about 900, which adds to the list of biologically active peptides which could be hormones. The high activity of the synthetic peptide in the rosette assay of the peptide synthesized on the basis of the amino acid sequence supports the specificity of its action. Its activity in various systems suggests that the natural peptide is involved in T cell differentiation. These include: O conversion in vitro and in vivo; enhancement of the generation of alloantigen reactive cytotoxic T cells in Tx mice; induction of suppressor T cells in NZB mice assayed using antibody production against polyvinyl pyrrolidone, enhancement of mitogen response in nude mice in vitro and ATx rats in vivo, and normalisation of the abnormally high level of autologous erythrocyte binding cells in ATx mice.

7/7/45 (Item 13 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

509662 EMBASE No: 93303849

South African psychiatrists' criteria for predicting dangerousness

Zabow T.; Cohen A.

Forensic Psychiatry Unit, University of Cape Town, Valkenberg Hospital,
Observatory, Cape Town South Africa

MED. LAW (South Africa), 1993, 12/3-5 (417-430) CODEN: MELAD ISSN:
0723-1393

LANGUAGES: English SUMMARY LANGUAGES: English

The prediction of dangerousness has emerged as one of the most controversial issues in forensic psychiatry. It is a value-laden and

ambiguous concept which has not been adequately defined and operationalized by the law or psychiatry. The validity and reliability of psychiatric predictions of dangerousness have been brought seriously into question in the past few decades. The Booysen Commission appointed in South Africa relates to the problem. An exploratory survey to ascertain the variables which South African psychiatrists perceive as influencing their decisions about dangerousness and to compare these findings with those of previous research was undertaken. One hundred and thirty-eight psychiatrists were included in a survey by means of a questionnaire. Findings that clinicians' decisions were significantly influenced by patients' criminal/violent history were consistent with previous research. The article further evaluates the responses to the questionnaire variables and the degree as well as a study of detained 'dangerous' patients to assess accuracy of evaluations in practice.

7/7/46 (Item 14 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

474828 EMBASE No: 76056790

Cellular cooperation during in vivo anti hapten antibody responses. III. The helper cell activity of activated thymocytes, of spleen cells treated with anti omega serum, and of spleen cells from anti thymocyte serum treated or adult thymectomized donors

Janeway C.A. Jr.

Lab. Immunol., NIAID, NIH, Bethesda, Md. 20014 USA

J. IMMUNOL. (USA), 1975, 114/4 (1408-1414) CODEN: JOIMA

LANGUAGES: ENGLISH

An adoptive secondary anti 2,4 dinitrophenyl (DNP) antibody response involving T B cell collaboration was studied. In particular, attempts were made to affect the unexpectedly steep log dose response curve obtained when graded numbers of helper cells are transferred to irradiated recipients given a fixed number of B cells (premium effect). A variety of means were used to alter helper cell activity, and this activity was then measured quantitatively, as was the ability of the helper cells present after these treatments to give a premium effect. It was shown that activated T cells are approximately twice as active as spleen cells in helper activity and give a comparable premium effect. Graded doses of anti theta serum plus complement markedly reduce the helper activity of spleen cells without affecting the premium effect given by the residual cells. Treatment of primed cell donors with limited doses of heterologous anti mouse thymocyte serum before transfer does not affect B cell activity, but readily inactivates helper cells, again without affecting the premium effect given by the residual cells. Adult thymectomy (ATx) of helper cell donors before priming with carrier led initially to increased helper activity relative to age matched control donors. This increase may reflect the loss of nonspecific suppressor T cells from spleens shortly after ATx. Late after ATx, there was about a 2 fold decrease in helper activity, probably reflecting a loss of helper cell precursors. At no time was there any change in the premium effect. In view of the failure of any of the techniques used to abolish the premium effect given by helper cells in this response, it seems likely that this premium effect is due to the cooperative interaction of 2 very similar types of mature T cell. Alternatively, the premium effect observed here may result from the interaction of 2 activities of a single type of T cell which is mediated by independent factors.

7/7/47 (Item 15 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1994 Elsevier Science B.V. All rts. reserv.

212068 EMBASE No: 75000046

Immunologic tolerance to a hapten. III. Induction of tolerance to trinitrophenyl in B cells in various differentiation states

The induction of tolerance to the TNP hapten in populations of B cells in various states of differentiation was studied with the free, reactive hapten trinitrobenzenesulfonic acid (TNBS). An adoptive transfer system was used in which tolerance was induced in an adoptively transferred B cell source before the addition of reconstituting thymus cells and immunization with hapten carrier. Bone marrow cells and cells from fetal liver were rendered tolerant to trinitrophenyl (TNP) in this manner, showing that tolerance can be induced in cells from the primary B cell precursor and stem cell source of both the adult and fetal mouse. Tolerance was induced in normal spleen cells after adoptive transfer and in spleen cells from adult thymectomized x-irradiated bone marrow reconstituted mouse (ATXBM) mice (B cell spleens) after transfer to adult thymectomized x-irradiated mouse (ATX) hosts. These experiments show the induction of tolerance to TNP in B cells in another state of differentiation (peripheral B cells). Tolerance could be induced in B cell sources depleted of T cells indicating that immunologic tolerance to TNP may be induced in B cells in the absence of T cells. The possibility of residual TNBS affecting the thymus cells used for reconstitution was ruled out by transferring fetal liver cells to a secondary host after tolerance induction in a primary adoptive host, and also by rendering ATXBM mice tolerant before the transfer of their spleen cells to an adoptive host. These maneuvers remove the B cells from the environment used for tolerance induction, and indicate that TNBS did not act upon reconstituting thymus cells. The data indicate that tolerance to TNP can be induced in early differentiation states of the B cell.

?ds

>>>Unrecognizable Command

?ds

Set	Items	Description
S1	46	AUTOCRINE(W)MOTILITY(W)FACTOR?. OR AUTOTAXIN OR ATX.(
S2	613	ATX
S3	649	S1 OR S2
S4	0	ANTIBO.D?
S5	1718826	ANTIBOD?
S6	73	S5 AND S3
S7	47	RD (unique items)

?rd s1

>>>Duplicate detection is not supported for File 398.

>>>Records from unsupported files will be retained in the RD set.

>>>Record 440:3418381 ignored; incomplete bibliographic data, not retained in RD set

...completed examining records

{ S8 26 RD S1 (unique items)

?s s8 not s7

{>>>Unrecognizable Command

?s s8 not s7

>>>Unrecognizable Command

?s s8 not s7

{ 26 S8

47 S7

S9 25 S8 NOT S7

?t s9/3/1-25

>>>Unrecognizable Command

?t s9/3/1-25

9/3/1 (Item 1 from file: 398)
 DIALOG(R)File 398:CHEMSEARCH(TM)
 (c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 150979-04-7

MOLECULAR FORMULA: Unknown

CA NAME(S):

HP=Peptide (9CI)

SB=(Val-Asn-Val-Ile-Ser-Gly-Pro-Ile-Phe-Asp-Tyr-Asp-Tyr-Asp-Gly-Leu-Xaa-Asp-Thr-Glu-Asp-Lys)

SYNONYMS: Autotaxin fragment (human)

9/3/2 (Item 2 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 150979-03-6

MOLECULAR FORMULA: Unknown

CA NAME(S):

HP=Peptide (9CI)

SB=(Thr-Phe-Pro-Asn-Leu-Tyr-Val-Xaa-Ala-Gln-Gly-Leu-Tyr-Trp-Ser)

SYNONYMS: Autotaxin fragment (human)

9/3/3 (Item 3 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 150979-02-5

MOLECULAR FORMULA: Unknown

CA NAME(S):

HP=Peptide (9CI)

SB=(Pro-Glu-Glu-Val-Thr-Xaa-Pro-Asn-Tyr-Leu)

SYNONYMS: Autotaxin fragment (human)

9/3/4 (Item 4 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 150979-01-4

MOLECULAR FORMULA: Unknown

CA NAME(S):

HP=Peptide (9CI)

SB=(Pro-Xaa-Leu-Asp-Val-Tyr-Lys)

SYNONYMS: Autotaxin fragment (human)

9/3/5 (Item 5 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 150236-72-9

MOLECULAR FORMULA: C74H99N13O19

RING SYSTEM DATA:

(03) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)

(04) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

CA NAME(S):

HP=L-Proline (9CI)

SB=L-tyrosylglycyl-L-phenylalanyl-L-leucyl-L-phenylalanyl-L-prolyl-L-prolyl-L-tyrosyl-L-leucyl-L-seryl-L-seryl-L-seryl-

SYNONYMS: Autotaxin fragment (human)

9/3/6 (Item 6 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147977-73-9

MOLECULAR FORMULA: C54H86N14O15

RING SYSTEM DATA:

(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)
(01) (nr=02; sr=5,6; ar=C4N.01-C6.01; fr=NC4.01-C6.01; ir=333-151-57)
CA NAME(S):
HP=L-Lysine (9CI)
SB=N2-(N-(N-(N-(N-(N-(N-(1-(N2-(N-glycylglycyl)-L-glutaminy)-L-prol
yl)-L-leucyl)-L-tryptophyl)-L-isoleucyl)-L-threonyl)-L-alanyl)-L-
threonyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/7 (Item 7 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-63-2
MOLECULAR FORMULA: C54H82N12O20
RING SYSTEM DATA:
(02) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)
(01) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Asparagine (9CI)
SB=N2-(N-(N-(1-(N-(N-(N-(N-(1-(N-L-seryl-L-tyrosyl)-L-prolyl)-L-.alp
ha.-glutamyl)-L-isoleucyl)-L-leucyl)-L-threonyl)-L-prolyl)-L-alan
yl)-L-.alpha.-aspartyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/8 (Item 8 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-62-1
MOLECULAR FORMULA: C63H96N16O14
RING SYSTEM DATA:
(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)
(01) (nr=01; sr=5; ar=C3N2.01; fr=NCNC2.01; ir=16-195-24)
(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Tyrosine (9CI)
SB=N-(N-(N-(N-(1-(N2-(N-(N-(N-(N-L-histidyl-L-leucyl)-L-leucyl)-L-ty
rosyl)glycyl)-L-arginyl)-L-prolyl)-L-alanyl)-L-valyl)-L-leucyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/9 (Item 9 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-61-0
MOLECULAR FORMULA: C26H40N6O9
RING SYSTEM DATA:
(01) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Threonine (9CI)
SB=N-(N-(N2-(N-L-tyrosyl-L-leucyl)-L-asparaginy)-L-alanyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/10 (Item 10 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-60-9
MOLECULAR FORMULA: C33H46N6O8
RING SYSTEM DATA:
(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

CA NAME(S):
HP=L-Phenylalanine (9CI)
SB=N-(N-(N2-(N-L-valyl-L-leucyl)-L-asparaginy)-L-tyrosyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/11 (Item 11 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-59-6
MOLECULAR FORMULA: C48H67N13O16
RING SYSTEM DATA:
(01) (nr=01; sr=5; ar=C3N2.01; fr=NCNC2.01; ir=16-195-24)
(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Serine (9CI)
SB=N-(N-(N-(N-(N2-(N-(N-(N-L-glutaminy)-L-tyrosyl)-L-leucyl)-L-histidyl)-L-glutaminy)-L-tyrosyl)glycyl)-L-seryl)-
SYNONYMS: Autotaxin fragment (human)

9/3/12 (Item 12 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-58-5
MOLECULAR FORMULA: C113H180N28O38
RING SYSTEM DATA:
(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)
(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Arginine (9CI)
SB=L-threonyl-L-.alpha.-glutamyl-L-phenylalanyl-L-leucyl-L-seryl-L-asparaginy)-L-tyrosyl-L-leucyl-L-threonyl-L-asparaginy)-L-valyl-L-.alpha.-aspartyl-L-.alpha.-aspartyl-L-isoleucyl-L-threonyl-L-leucyl-L-valyl-L-prolylglycyl-L-threonyl-L-leucylglycyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/13 (Item 13 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-57-4
MOLECULAR FORMULA: C68H101N17O21
RING SYSTEM DATA:
(01) (nr=01; sr=5; ar=C3N2.01; fr=NCNC2.01; ir=16-195-24)
(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)
CA NAME(S):
HP=L-Arginine (9CI)
SB=N2-(N-(N-(N-(N-(N-(N-(N-(N-(N-L-.alpha.-aspartyl-L-isoleucyl)-L-.alpha.-glutamyl)-L-histidyl)-L-leucyl)-L-threonyl)-L-seryl)-L-leucyl)-L-.alpha.-aspartyl)-L-phenylalanyl)-L-phenylalanyl)-
SYNONYMS: Autotaxin fragment (human)

9/3/14 (Item 14 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-56-3
MOLECULAR FORMULA: C82H123N19O23S
RING SYSTEM DATA:
(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)
(03) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

CA NAME(S):

HP=L-Lysine (9CI)

SB=L-valyl-L-asparaginyL-L-seryl-L-methionyl-L-glutaminyL-L-threonyL
-L-valyl-L-phenylalanyl-L-valylglycyl-L-tyrosylglycyl-L-prolyl-L-
threonyL-L-phenylalanyl-

SYNONYMS: Autotaxin fragment (human)

9/3/15 (Item 15 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-55-2

MOLECULAR FORMULA: C56H80N12O17

RING SYSTEM DATA:

(02) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)

(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

CA NAME(S):

HP=L-Tyrosine (9CI)

SB=N-(N-(N-(N2-(N-(N-(1-(1-L-seryl-L-prolyl)-L-prolyl)-L-phenyla
lanyl)-L-.alpha.-glutamyl)-L-asparaginyL)-L-isoleucyl)-L-asparagi
nyl)-L-leucyl)-

SYNONYMS: Autotaxin fragment (human)

9/3/16 (Item 16 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-54-1

MOLECULAR FORMULA: C53H73N11O17

RING SYSTEM DATA:

(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)

(01) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

(01) (nr=02; sr=5,6; ar=C4N.01-C6.01; fr=NC4.01-C6.01; ir=333-151-57)

CA NAME(S):

HP=L-Isoleucine (9CI)

SB=N-(N-(N-(N2-(N-(1-(N-(N-L-tyrosyl-L-.alpha.-aspartyl)-L-valyl)-L-
prolyl)-L-tryptophyl)-L-asparaginyL)-L-.alpha.-glutamyl)-L-threon
yl)-

SYNONYMS: Autotaxin fragment (human)

9/3/17 (Item 17 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-53-0

MOLECULAR FORMULA: C21H36N6O10

CA NAME(S):

HP=L-Serine (9CI)

SB=N-(N-(N-(N-L-glutaminyL-L-alanyl)-L-.alpha.-glutamyl)-L-valyl)-

SYNONYMS: Autotaxin fragment (human)

9/3/18 (Item 18 from file: 398)

DIALOG(R)File 398:CHEMSEARCH(TM)

(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-52-9

MOLECULAR FORMULA: C32H44N6O7

RING SYSTEM DATA:

(01) (nr=01; sr=5; ar=C4N.01; fr=NC4.01; ir=16-136-1)

(02) (nr=01; sr=6; ar= fr=C6.01; ir=46-150-18)

CA NAME(S):

HP=L-Lysine (9CI)

SB=N2-(N-(N-(1-L-tyrosyl-L-prolyl)-L-alanyl)-L-phenylalanyl)-

SYNONYMS: Autotaxin fragment (human)

9/3/19 (Item 19 from file: 398)
DIALOG(R)File 398:CHEMSEARCH(TM)
(c) 1994 Amer.Chem.Soc. All rts. reserv.

CAS REGISTRY NUMBER: 147960-51-8

MOLECULAR FORMULA: C32H44N10O8

RING SYSTEM DATA:

(01) (nr=01; sr=5; ar=C3N2.01; fr=NCNC2.01; ir=16-195-24)

(01) (nr=02; sr=5,6; ar=C4N.01-C6.01; fr=NC4.01-C6.01; ir=333-151-57)

CA NAME(S):

HP=L-Asparagine (9CI)

SB=N2-(N-(N-(N-(N-L-tryptophyl-L-histidyl)-L-valyl)-L-alanyl)-L-alan
yl)-

SYNONYMS: Autotaxin fragment (human)

9/3/20 (Item 1 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

11147913 BIOSIS Number: 97347913

Partial cDNA cloning of the autocrine cell motility factor, autotaxin
(ATX)

Murata J; Arestad A; Licotta L A; Stracke M L

Natl. Inst. Health, Natl. Cancer Inst., Bethesda, MD 20892, USA

FASEB Journal 8 (7). 1994. A1445.

Full Journal Title: 85th Annual Meeting of the American Society for
Biochemistry and Molecular Biology, Washington, D.C., USA, May 21-25, 1994.

FASEB Journal

ISSN: 0892-6638

Language: ENGLISH

9/3/21 (Item 2 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

11147796 BIOSIS Number: 97347796

Glycosylation of the tumor cell motility factor, autotaxin, is not
required for activity

Stracke M L; Levine M; Arestad A; Liotta L A

Natl. Inst. Health, Natl. Cancer Inst., Bethesda, MD 20892, USA

FASEB Journal 8 (7). 1994. A1424.

Full Journal Title: 85th Annual Meeting of the American Society for
Biochemistry and Molecular Biology, Washington, D.C., USA, May 21-25, 1994.

FASEB Journal

ISSN: 0892-6638

Language: ENGLISH

9/3/22 (Item 3 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

11053808 BIOSIS Number: 97253808

Characterization of a new scatter factor (CL4-SA) produced by rat mammary
tumor cells

Kopdag H; Hoelzel F; Scherдин U

Inst. Physiol. Chem., Univ. Hosp., Hamburg-Eppendorf, GER

Journal of Cancer Research and Clinical Oncology 120 (SUPPL.). 1994.

R114.

Full Journal Title: 21st National Cancer Congress of the German Cancer

9/3/23 (Item 4 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

10406739 BIOSIS Number: 96006739
GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR INDUCES HUMAN
MELANOMA-CELL MIGRATION
KOHN E C; HOLLISTER G H; DIPERSIO J D; WAHL S; LIOTTA L A; SCHIFFMANN E
BLDG. 10, ROOM 2A33, NATL. CANCER INST., BETHESDA, MD 20892, USA.
INT J CANCER 53 (6). 1993. 968-972. CODEN: IJCNA
Full Journal Title: International Journal of Cancer
Language: ENGLISH

9/3/24 (Item 5 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1994 BIOSIS. All rts. reserv.

9093687 BIOSIS Number: 93078687
IDENTIFICATION PURIFICATION AND PARTIAL SEQUENCE ANALYSIS OF AUTOTAXIN A
NOVEL MOTILITY-STIMULATING PROTEIN
STRACKE M L; KRUTZSCH H C; UNSWORTH E J; ARESTAD A; CIOCE V; SCHIFFMANN E
; LIOTTA L A
LAB. PATHOL., NATIONAL CANCER INST., NATIONAL INST. HEALTH, BUILDING 10,
ROOM 2A33, BETHESDA, MD. 20892.
J BIOL CHEM 267 (4). 1992. 2524-2529. CODEN: JBCHA
Full Journal Title: Journal of Biological Chemistry
Language: ENGLISH

9/3/25 (Item 1 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1994 Dialog Info.Svcs. All rts. reserv.

08903820 94218820
The role of autotaxin and other motility stimulating factors in the
regulation of tumor cell motility.
Stracke M; Liotta LA; Schiffmann E
Laboratory of Pathology, National Cancer Institute, National Institutes
of Health, Bethesda, MD 20892.
Symp Soc Exp Biol (ENGLAND) 1993, 47 p197-214, ISSN 0081-1386
Journal Code: VGF
Languages: ENGLISH
Document type: JOURNAL ARTICLE
?e au=~~6f6acke~~em
>>>Unrecognizable Command
?e au=stracke, m

Ref	Items	Index-term
E1	3	AU=STRACKE, K. J.
E2	1	AU=STRACKE, K.H.
E3	0	*AU=STRACKE, M
E4	7	AU=STRACKE, M.
E5	1	AU=STRACKE, MARICA
E6	9	AU=STRACKE, MARKUS
E7	2	AU=STRACKE, MARY
E8	11	AU=STRACKE, MARY L.
E9	2	AU=STRACKE, MICHAEL
E10	5	AU=STRACKE. R.

E11 1 AU=STRACKE, ROBERT J.
E12 2 AU=STRACKE, ROLAND

Enter P or PAGE for more

?{s s

?s e4 or e7

>>>One or more prefixes are unsupported

>>> or undefined in one or more files.

7 AU=STRACKE, M.

2 AU=STRACKE, MARY

S10 9 AU="STRACKE, M." OR AU="STRACKE, MARY"

?fd

>>>Unrecognizable Command

?ds

Set	Items	Description
S1	46	AUTOCRINE(W)MOTILITY(W)FACTOR?. OR AUTOTAXIN OR ATX.(
S2	613	ATX
S3	649	S1 OR S2
S4	0	ANTIBO.D?
S5	1718826	ANTIBOD?
S6	73	S5 AND S3
S7	47	RD (unique items)
S8	26	RD S1 (unique items)
S9	25	S8 NOT S7
S10	9	AU="STRACKE, M." OR AU="STRACKE, MARY"

?fd s10

>>>Duplicate detection is not supported for File 398.

>>>Records from unsupported files will be retained in the RD set.

...completed examining records

S11 5 .RD S10 (unique items)

?s s11 not s8

5 S11

26 S8

S12 4 S11 NOT S8

?s s12 not s7

4 S12

47 S7

S13 4 S12 NOT S7

{?t {s13/

{

>>>Unrecognizable Command

?t s13/3/1-4

13/3/1 (Item 1 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

121079613 CA: 121(7)79613a JOURNAL

The role of autotaxin and other motility stimulating factors in the regulation of tumor cell motility

AUTHOR(S): Stracke, Mary; Liotta, Lance A.; Schiffmann, Elliott

LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA

JOURNAL: Symp. Soc. Exp. Biol. DATE: 1993 VOLUME: 47 NUMBER: CELL

BEHAVIOUR PAGES: 197-214 CODEN: SSEBA9 ISSN: 0081-1386 LANGUAGE: English

13/3/2 (Item 2 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

110190136 CA: 110(21)190136j JOURNAL

Biochemical mechanisms of tumor invasion and metastases

AUTHOR(S): Liotta, L. A.; Wewer, U.; Rao, N. C.; Schiffmann, E.; Stracke,

M.; Guirguis, R.; Thorgerirsson, U.; Muschel, R.; Sobel, M.
LOCATION: Lab. Pathol., NIH, Bethesda, MD, 20892, USA
JOURNAL: Adv. Exp. Med. Biol. DATE: 1988 VOLUME: 233 NUMBER: Cancer
Metastasis: Biol. Biochem. Mech. Clin. Aspects PAGES: 161-9 CODEN:
AEMBAP ISSN: 0065-2598 LANGUAGE: English

13/3/3 (Item 1 from file: 6)
DIALOG(R)File 6:NTIS
Comp. & distr. 1994 NTIS, US Dept of Commerce. All rts. reserv.

1491515 NTIS Accession Number: TIB/A90-81562/XAB
Belastungs- und Beulversuche an axialsymmetrisch belasteten
Rotationsschalen aus Metall im elastisch-plastischen Bereich zur
Ueberpruefung nichtlinearer Rechenprogramme. (Load and buckling tests and
axisymmetrically loaded metal shells of revolution in the elastic-plastic
range for the verification of nonlinear computer programs)
Stracke, M. ; Duesing, H. ; Krysik, R. ; Schmidt, H.
Gesamthochschule Essen (Germany, F.R.). Fachbereich 10 - Bauwesen.
Corp. Source Codes: 049146012; 9900856
Sep 86 293p
Languages: German
Journal Announcement: GRAI9101
In German. Forschungsberichte aus dem Fachbereich Bauwesen, Universitaet
- Gesamthochschule Essen, no. 38.
NTIS Prices: PC E07

13/3/4 (Item 2 from file: 6)
DIALOG(R)File 6:NTIS
Comp. & distr. 1994 NTIS, US Dept of Commerce. All rts. reserv.

1308743 NTIS Accession Number: TIB/A88-80048/XAB
Beulversuche an laengsnahtgeschweissten staehlernen Kreiszyinderschalen
unter Aussendruck im elastisch-plastischen Bereich. (Buckling tests on
longitudinally welded cylindrical steel shells under external pressure in
the elasto-plastic range)
Stracke, M. ; Schmidt, H.
Gesamthochschule Essen (Germany, F.R.). Fachbereich 10 - Bauwesen.
Corp. Source Codes: 049146012
Dec 84 10p.
Languages: German
Journal Announcement: GRAI8810
In German. Forschungsberichte aus dem Fachbereich Bauwesen, Universitaet
Essen, Gesamthochschule, no. 28.
NTIS Prices: PC E07

?e a@#}0ptta, 1
>>>Unrecognizable Command
?e au€#iota, 1

Ref	Items	Index-term
E1	0	*AU.=LIOTTA, L
E2	1	AU=A CRIEPI, N. S.
E3	1	AU=A. B. SIDDIQUE
E4	1	AU=A. FENG
E5	1	AU=A. WULF, WM.
E6	1	AU=A-DAVIDSON, REVATHI
E7	3	AU=A-GRIVAS, D.
E8	1	AU=A-GRIVAS, DIMITRI
E9	1	AU=A-H TAYEL, M.
E10	1	AU=A-KWIE JAP TJOEN SAN, E. R.
E11	3	AU=A-MONEIM, M. T.
E12	1	AU=A, SHAFER

Enter P or PAGE for more
?e au=liotta, 1

Ref	Items	Index-term
E1	1	AU=LIOTTA, JUDITH J.
E2	1	AU=LIOTTA, K.
E3	0	*AU=LIOTTA, L
E4	12	AU=LIOTTA, L.
E5	72	AU=LIOTTA, L. A.
E6	1	AU=LIOTTA, L. J.
E7	8	AU=LIOTTA, LANCE
E8	164	AU=LIOTTA, LANCE A.
E9	1	AU=LIOTTA, LANCE ALLEN
E10	8	AU=LIOTTA, LOUIS J.
E11	1	AU=LIOTTA, LOUIS JAMES
E12	1	AU=LIOTTA, MARIO

Enter P or PAGE for more

?s e4 or e5 or e7 or e8 or e9

>>>Term "O.R" in invalid position

?s e4 or e5 or e7 or e8 or e9

>>>One or more prefixes are unsupported

>>> or undefined in one or more files.

12 AU=LIOTTA, L.

72 AU=LIOTTA, L. A.

8 AU=LIOTTA, LANCE

164 AU=LIOTTA, LANCE A.

1 AU=LIOTTA, LANCE ALLEN

S14 257 AU="LIOTTA, L." OR AU="LIOTTA, L. A." OR AU="LIOTTA, LANCE" OR AU="LIOTTA, LANCE A." OR AU="LIOTTA, LANCE ALLEN"

?rd

>>>Duplicate detection is not supported for File 398.

>>>Records from unsupported files will be retained in the RD set.

...examined 50 records (50)

...examined 50 records (100)

...examined 50 records (150)

...examined 50 records (200)

...examined 50 records (250)

...completed examining records

S15 247 RD (unique items)

?ds

Set	Items	Description
S1	46	AUTOCRINE(W)MOTILITY(W)FACTOR?. OR AUTOTAXIN OR ATX.(
S2	613	ATX
S3	649	S1 OR S2
S4	0	ANTIBO.D?
S5	1718826	ANTIBOD?
S6	73	S5 AND S3
S7	47	RD (unique items)
S8	26	RD S1 (unique items)
S9	25	S8 NOT S7
S10	9	AU="STRACKE, M." OR AU="STRACKE, MARY"
S11	5	.RD S10 (unique items)
S12	4	S11 NOT S8
S13	4	S12 NOT S7
S14	257	AU="LIOTTA, L." OR AU="LIOTTA, L. A." OR AU="LIOTTA, LANCE" OR AU="LIOTTA, LANCE A." OR AU="LIOTTA, LANCE ALLEN"
S15	247	RD (unique items)
?s s15 not s8		
	247	S15
	26	S8
S16	246	S15 NOT S8
?s s16 not s7		
	246	S16
	47	S7

S17 246 S16 NOT S7
?s s17 not s13
246 S17
4 S13
S18 244 S17 NOT S13
?s s18 and s3
244 S18
649 S3
S19 1 S18 AND S3

?t s19/7/1

{
19/7/1 (Item 1 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

118144592 CA: 118(15)144592x JOURNAL
Identification, purification, and partial sequence analysis of autotaxin,
a novel motility-stimulating protein
AUTHOR(S): Stracke, Mary L.; Krutzsch, Henry C.; Unsworth, Edward J.;
Arestad, Anders; Cioce, Vittoria; Schiffmann, Elliott; Liotta, Lance A.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: J. Biol. Chem. DATE: 1992 VOLUME: 267 NUMBER: 4 PAGES:
2524-9 CODEN: JBCHA3 ISSN: 0021-9258 LANGUAGE: English
SECTION:
CA214001 Mammalian Pathological Biochemistry
IDENTIFIERS: autotaxin melanoma autocrine motility factor
DESCRIPTORS:
Nomenclature, new natural products...
autotaxin (human protein)
Melanoma...
autotaxin as autocrine motility factor from human A2058 cell line of,
identification and purifn. and partial amino acid sequence of
Proteins, specific or class...
autotaxins, identification and purifn. and partial amino acid sequence
of, as autocrine motility factor from A2058 melanoma cell line of human
G proteins (guanine nucleotide-binding proteins), Gi (adenylate
cyclase-inhibiting)...
in autotaxin-mediated human A2058 melanoma cell motility stimulation
Protein sequences...
of autotaxin peptides, of human
CAS REGISTRY NUMBERS:
146589-06-2 amino acid sequence of
?e au=schiffmann, e
>>>Unrecognizable Command
?e au=sbhiffmann, e

Ref	Items	Index-term
E1	1	AU=SCHIFFMANN, DEITMAR
E2	24	AU=SCHIFFMANN, DIETMAR
E3	0	*AU=SCHIFFMANN, E
E4	21	AU=SCHIFFMANN, E.
E5	1	AU=SCHIFFMANN, ELIOTT
E6	10	AU=SCHIFFMANN, ELLIOT
E7	46	AU=SCHIFFMANN, ELLIOTT
E8	1	AU=SCHIFFMANN, F.
E9	1	AU=SCHIFFMANN, GENEVIEVE N.
E10	3	AU=SCHIFFMANN, H.
E11	2	AU=SCHIFFMANN, HEINRICH
E12	2	AU=SCHIFFMANN, I.

Enter P or PAGE for more
?e e4 6ore5 or e6 or e7

Ref	Items	Index-term
E1	1	E399-81
E2	8137	E4

E3		0	*E4 .OR E5 OR E6 OR E7
E4		37	E4 //LEUKOTRIENE
E5		438	E4 DURING ANAPHYLAXIS AND INFLAMM//SYNTHESIS,
E6		26	E4 GENE
E7		12	E4 GENE-PRODUCTS
E8		2	E4 PROMOTER
E9		2	E4 PROTEIN
E10		1	E4 PROTEIN, TOMATO
E11		4	E4 PROTEINS
E12		35	E4 PROTEINS //ADENOVIRUS

Enter P or PAGE for more

?fse4:e7

S20 513 "E4 //LEUKOTRIENE": "E4 GENE-PRODUCTS"

?e au=schiffmann; e

Ref	Items	Index-term
E1	7	AU=SCHIFFMANN Z
E2	1	AU=SCHIFFMANN Z.
E3	0	*AU=SCHIFFMANN., E
E4	1	AU=SCHIFFMANN-NADEL
E5	86	AU=SCHIFFMANN-NADEL M
E6	2	AU=SCHIFFMANN-NADEL, M.
E7	16	AU=SCHIFFMANN-NADEL, MINA
E8	3	AU=SCHIFFMANN-WYTTEBACH E
E9	1	AU=SCHIFFMANN, A.
E10	2	AU=SCHIFFMANN, A. P.
E11	1	AU=SCHIFFMANN, ALOIS
E12	1	AU=SCHIFFMANN, ANNAMARIE

Enter P or PAGE for more

?e{au=schiffmann, e

>>>Unrecognizable Command

?e au=schiffmann, e

Ref	Items	Index-term
E1	1	AU=SCHIFFMANN, DEITMAR
E2	24	AU=SCHIFFMANN, DIETMAR
E3	0	*AU=SCHIFFMANN, E
E4	21	AU=SCHIFFMANN, E.
E5	1	AU=SCHIFFMANN, ELIOTT
E6	10	AU=SCHIFFMANN, ELLIOTT
E7	46	AU=SCHIFFMANN, ELLIOTT
E8	1	AU=SCHIFFMANN, F.
E9	1	AU=SCHIFFMANN, GENEVIEVE N.
E10	3	AU=SCHIFFMANN, H.
E11	2	AU=SCHIFFMANN, HEINRICH
E12	2	AU=SCHIFFMANN, I.

Enter P or PAGE for more

?f e4:e7

>>>One or more prefixes are unsupported

>>> or undefined in one or more files.

S21 78 AU="SCHIFFMANN, E.":AU="SCHIFFMANN, ELLIOTT"

{?rd

>>>Unrecognizable Command

?rd

>>>Duplicate detection is not supported for File 398.

{

>>>Records from unsupported files will be retained in the RD set.

...examined 50 records (50)

...completed examining records

S22 72 RD (unique items)

?s s22 not s7

<>>>Unrecognizable Command

?fss22 not s7

>>>Unrecognizable Command

?s s22 not s7

72 S22

47 S7

S23 71 S22 NOT S7

?s s23 not s8

71 S23

26 S8

S24 71 S23 NOT S8

?s s24 not s19

71 S24

1 S19

S25 70 S24 NOT S19

?s s25 not s13

70 S25

4 S13

S26 68 S25 NOT S13

{?{

?s s26 and s3

0 S.226

649 S3

S27 0 S.226 AND S3

?{

>>>Unrecognizable Command

?s s26 and s3

68 S26

649 S3

S28 0 S26 AND S3

?e au=krutzsch, h

Ref Items Index-term

E1 1 AU=KRUTZSCH, E.

E2 0 *AU=KRUTZSCH, H

E3 6 AU=KRUTZSCH, H.

E4 7 AU=KRUTZSCH, H. C.

E5 17 AU=KRUTZSCH, HENRY

E6 50 AU=KRUTZSCH, HENRY C.

E7 1 AU=KRUTZSCH, HENTRY C.

E8 3 AU=KRUTZSCH, JOHANNES

E9 1 AU=KRUTZSCH, P. H.

E10 1 AU=KRUTZSCH, PH.

E11 4 AU=KRUTZSCH, PHILIP H.

E12 1 AU=KRUTZSCH, PHILIP H,

Enter P or PAGE for more

?s e3:e7

>>>Invalid term: E.{7

?

?s e3:e7

>>>One or more prefixes are unsupported

>>> or undefined in one or more files.

S29 81 AU="KRUTZSCH, H.":AU="KRUTZSCH, HENTRY C."

?fd

{>>>Duplicate detection is not supported for File 398.

>>>Records from unsupported files will be retained in the RD set.

...examined 50 records (50)

...completed examining records

S30 78 .RD (unique items)

?s s30 not (s7 or s8 or s13 or s19)

>>>Unrecognizable Command

?s s30 not (s7 or s8 or s13 or s19)

{ 78 S30

47 S7

26 S8

4 S13

1 S19
S31 77 S30 NOT (S7 OR S8 OR S13 OR S19)
?t s31/3/1-20
{>>>Unrecognizable Command
?t s31/3/1-20
>>>Unrecognizable Command
?t s31/3/1-20

31/3/1 (Item 1 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

121079330 CA: 121(7)79330f JOURNAL
High-affinity .alpha.-thrombin binding to platelet glycoprotein
Ib.alpha.: identification of two binding domains
AUTHOR(S): Gralnick, Harvey R.; Williams, Sybil; McKeown, Laurie P.;
Hansmann, Kristin; Feinton, John W., II; Krutzsch, Henry
LOCATION: Hematology Service, National Institutes Health, Bethesda, MD,
20892, USA
JOURNAL: Proc. Natl. Acad. Sci. U. S. A. DATE: 1994 VOLUME: 91
NUMBER: 14 PAGES: 6334-8 CODEN: PNASA6 ISSN: 0027-8424 LANGUAGE:
English

31/3/2 (Item 2 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

120262913 CA: 120(21)262913s JOURNAL
A sequence-specific, single-strand binding protein activates the far
upstream element of c-myc and defines a new DNA-binding motif
AUTHOR(S): Duncan, Robert; Bazar, Leonard; Michelotti, Greg; Tomonaga,
Takeshi; Krutzsch, Henry; Avigan, mark; Levens, David
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Genes Dev. DATE: 1994 VOLUME: 8 NUMBER: 4 PAGES: 465-80
CODEN: GEDEEP ISSN: 0890-9369 LANGUAGE: English

31/3/3 (Item 3 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

120208113 CA: 120(17)208113t JOURNAL
Apolipoprotein E: a potent inhibitor of endothelial and tumor cell
proliferation
AUTHOR(S): Vogel, Tikva; Guo, Nenghua; Guy, Rachel; Drezlich, Nina;
Krutzsch, Henry C.; Blake, Diane A.; Panet, Amos; Roberts, David D.
LOCATION: Lab. Pathol., Natl. Inst. Health, Bethesda, MD, 20892, USA
JOURNAL: J. Cell. Biochem. DATE: 1994 VOLUME: 54 NUMBER: 3 PAGES:
299-308 CODEN: JCEBD5 ISSN: 0730-2312 LANGUAGE: English

31/3/4 (Item 4 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119262220 CA: 119(25)262220t JOURNAL
Modulation of endothelial cell proliferation, adhesion, and motility by
recombinant heparin-binding domain and synthetic peptides from the type I
repeats of thrombospondin
AUTHOR(S): Vogel, Tikva; Guo, Neng Hua; Krutzsch, Henry C.; Blake, Diane
A.; Hartman, Jacob; Mendelovitz, Simona; Panet, Amos; Roberts, David D.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: J. Cell. Biochem. DATE: 1993 VOLUME: 53 NUMBER: 1 PAGES:
74-84 CODEN: JCEBD5 ISSN: 0730-2312 LANGUAGE: English

31/3/5 (Item 5 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119198234 CA: 119(19)198234w JOURNAL
The purification and characterization of an extremely thermostable
.alpha.-amylase from the hyperthermophilic archaebacterium *Pyrococcus
furiosus*
AUTHOR(S): Laderman, Kenneth A.; Davis, Bradley R.; Krutzsch, Henry C.;
Lewis, Marc S.; Griko, Y. V.; Privalov, Peter L.; Anfinsen, Christian B.
LOCATION: Dep. Biol., Johns Hopkins Univ., Baltimore, MD, 21218, USA
JOURNAL: J. Biol. Chem. DATE: 1993 VOLUME: 268 NUMBER: 32 PAGES:
24394-401 CODEN: JBCHA3 ISSN: 0021-9258 LANGUAGE: English

31/3/6 (Item 6 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119175324 CA: 119(17)175324m JOURNAL
Specific binding of heterogeneous ribonucleoprotein particle protein K to
the human c-myc promoter, in vitro
AUTHOR(S): Takimoto, Masato; Tomonaga, Takeshi; Matunis, Michael; Avigan,
Mark; Krutzsch, Henry; Dreyfuss, Gideon; Levens, David
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: J. Biol. Chem. DATE: 1993 VOLUME: 268 NUMBER: 24 PAGES:
18249-58 CODEN: JBCHA3 ISSN: 0021-9258 LANGUAGE: English

31/3/7 (Item 7 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119023356 CA: 119(3)23356k CONFERENCE PROCEEDING
TIMP-2: identification and characterization of a new member of the
metalloproteinase inhibitor family
AUTHOR(S): Stetler-Stevenson, William G.; Krutzsch, Henry C.; Liotta,
Lance A.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Matrix Metalloproteinases Inhib., Proc. Matrix Metalloproteinase
Conf. EDITOR: Birkedal-Hansen, Henning (Ed), DATE: 1992 PAGES: 299-306
CODEN: 58AQAX LANGUAGE: English MEETING DATE: 890000 PUBLISHER:
Fischer, Stuttgart, Germany

31/3/8 (Item 8 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

119005737 CA: 119(1)5737f JOURNAL
Inhibition of fibronectin binding and fibronectin-mediated cell adhesion
to collagen by a peptide from the second type I repeat of thrombospondin
AUTHOR(S): Sipes, John M.; Guo, Neng Hua; Negre, Eric; Vogel, Tikva;
Krutzsch, Henry C.; Roberts, David D.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: J. Cell Biol. DATE: 1993 VOLUME: 121 NUMBER: 2 PAGES: 469-77
CODEN: JCLBA3 ISSN: 0021-9525 LANGUAGE: English

31/3/9 (Item 9 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

118250600 CA: 118(25)250600h JOURNAL
BAC-1: a mitogen-induced nuclear protein tyrosine phosphatase

AUTHOR(S): Rohan, Patricia J.; Davis, Paula; Moskaluk, Christopher A.;
Kearns, Mary; Krutzsch, Henry; Siebenlist, Ulrich; Kelly, Kathleen
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Science (Washington, D. C., 1983-) DATE: 1993 VOLUME: 259
NUMBER: 5102 PAGES: 1763-6 CODEN: SCIEAS ISSN: 0036-8075 LANGUAGE:
English

31/3/10 (Item 10 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

118142727 CA: 118(15)142727w JOURNAL
N-isopropylidoacetamide in the reduction and alkylation of proteins:
Use in microsequence analysis
AUTHOR(S): Krutzsch, Henry C.; Inman, John K.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, USA
JOURNAL: Anal. Biochem. DATE: 1993 VOLUME: 209 NUMBER: 1 PAGES:
109-16 CODEN: ANBCA2 ISSN: 0003-2697 LANGUAGE: English

31/3/11 (Item 11 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

118140848 CA: 118(15)140848n JOURNAL
Cloning and characterization of a novel human cDNA that has DNA
similarity to the conserved region of the collagenase gene family
AUTHOR(S): Templeton, Nancy Smyth; Rodgers, Lisa A.; Levy, Anna T.; Ting,
Kai Li; Krutzsch, Henry C.; Liotta, Lance A.; Stetler-Stevenson, William
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Genomics DATE: 1992 VOLUME: 12 NUMBER: 1 PAGES: 175-6
CODEN: GNMCEP ISSN: 0888-7543 LANGUAGE: English

31/3/12 (Item 12 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

117209613 CA: 117(21)209613f JOURNAL
A monomeric von Willebrand factor fragment, Leu-504-Ser-728, inhibits von
Willebrand factor interaction with glycoprotein Ib-IX
AUTHOR(S): Gralnick, Harvey R.; Williams, Sybil; McKeown, Laurie; Kramer,
Wendy; Krutzsch, Henry; Gorecki, Marian; Pinet, Amos; Garfinkel, Leonard I.
LOCATION: Hematol. Serv., Natl. Inst. Health, Bethesda, MD, 20892, USA
JOURNAL: Proc. Natl. Acad. Sci. U. S. A. DATE: 1992 VOLUME: 89
NUMBER: 17 PAGES: 7800-4 CODEN: PNASA6 ISSN: 0027-8424 LANGUAGE:
English

31/3/13 (Item 13 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

117189539 CA: 117(19)189539f JOURNAL
Heparin-binding peptides from the type I repeats of thrombospondin.
Structural requirements for heparin binding and promotion of melanoma cell
adhesion and chemotaxis
AUTHOR(S): Guo, Neng Hua; Krutzsch, Henry C.; Negre, Eric; Zabrenetzky,
Vivian S.; Roberts, David D.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: J. Biol. Chem. DATE: 1992 VOLUME: 267 NUMBER: 27 PAGES:
19349-55 CODEN: JBCHA3 ISSN: 0021-9258 LANGUAGE: English

31/3/14 (Item 14 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

117106652 CA: 117(11)106652p JOURNAL

Interactions of a laminin-binding peptide from a 33-kDa protein related to the 67-kDa laminin receptor with laminin and melanoma cells are heparin-dependent

AUTHOR(S): Guo, Neng Hua; Krutzsch, Henry C.; Vogel, Tikva; Roberts, David D.

LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA

JOURNAL: J. Biol. Chem. DATE: 1992 VOLUME: 267 NUMBER: 25 PAGES: 17743-7 CODEN: JBCHA3 ISSN: 0021-9258 LANGUAGE: English

31/3/15 (Item 15 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

117085872 CA: 117(9)85872t JOURNAL

A proteolytically sensitive region common to several rat liver cytochromes P450: effect of cleavage on substrate binding

AUTHOR(S): Tsokos, Dimitris C.; Omata, Yoshiaki; Robinson, Richard C.; Krutzsch, Henry C.; Gelboin, Harry V.; Friedman, Fred K.

LOCATION: Lab. Mol. Carcinog., Natl. Cancer Inst., Bethesda, MD, 20892, USA

JOURNAL: Biochemistry DATE: 1992 VOLUME: 31 NUMBER: 31 PAGES: 7155-9 CODEN: BICHAW ISSN: 0006-2960 LANGUAGE: English

31/3/16 (Item 16 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

117063782 CA: 117(7)63782m PATENT

Human megakaryocyte colony-stimulating factor (hMeg-CSF) protein and methods

INVENTOR(AUTHOR): Murphy, Martin J.; Parchment, Ralph E.; Erickson-Miller, Connie L.; Dai, Wei; Zhang, Zhao Geng; Liotta, Lance A.; Krutzsch, Henry

LOCATION: USA

ASSIGNEE: Hipple Cancer Research Center

PATENT: PCT International ; WO 9200319 A1 DATE: 920109

APPLICATION: WO 91US4698 (910702) *US 547573 (900702)

PAGES: 86 pp. CODEN: PIXXD2 LANGUAGE: English CLASS: C07K-003/28A; C07K-017/00B DESIGNATED COUNTRIES: AU; CA; FI; JP; KR; NO

DESIGNATED REGIONAL: AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU; NL; SE

31/3/17 (Item 17 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

116232948 CA: 116(23)232948g JOURNAL

Heparin- and sulfatide-binding peptides from the type I repeats of human thrombospondin promote melanoma cell adhesion

AUTHOR(S): Guo, Neng Hua; Krutzsch, Henry C.; Negre, Eric; Vogel, Tikva; Blake, Diane A.; Roberts, David D.

LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA

JOURNAL: Proc. Natl. Acad. Sci. U. S. A. DATE: 1992 VOLUME: 89 NUMBER: 7 PAGES: 3040-4 CODEN: PNASA6 ISSN: 0027-8424 LANGUAGE: English

31/3/18 (Item 18 from file: 399)

DIALOG(R)File 399:CA Search(R)

(c) 1994 American Chemical Society. All rts. reserv.

116101599 CA: 116(11)101599w JOURNAL
Higher-order complex formation between the 72-kilodalton type IV
collagenase and tissue inhibitor of metalloproteinases-2
AUTHOR(S): Kleiner, David E., Jr.; Unsworth, Edward J.; Krutzsch, Henry
C.; Stetler-Stevenson, William G.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Biochemistry DATE: 1992 VOLUME: 31 NUMBER: 6 PAGES: 1665-72
CODEN: BICHAW ISSN: 0006-2960 LANGUAGE: English

31/3/19 (Item 19 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

115227417 CA: 115(21)227417h JOURNAL
Reducing chemical background noise in automated protein sequencers
AUTHOR(S): Fransworth, Vince; Carson, Wulf; Krutzsch, Henry
LOCATION: Porton Instrum., Tarzana, CA, 91356, USA
JOURNAL: Pept. Res. DATE: 1991 VOLUME: 4 NUMBER: 4 PAGES: 245-51
CODEN: PEREEO ISSN: 1040-5704 LANGUAGE: English

31/3/20 (Item 20 from file: 399)
DIALOG(R)File 399:CA Search(R)
(c) 1994 American Chemical Society. All rts. reserv.

115046587 CA: 115(5)46587t JOURNAL
Biosynthesis of the 67 kDa high affinity laminin receptor
AUTHOR(S): Castronovo, Vincent; Claysmith, Anne P.; Barker, Karen T.;
Cioce, Vittoria; Krutzsch, Henry C.; Sobel, Mark E.
LOCATION: Lab. Pathol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
JOURNAL: Biochem. Biophys. Res. Commun. DATE: 1991 VOLUME: 177
NUMBER: 1 PAGES: 177-83 CODEN: BBRC A9 ISSN: 0006-291X LANGUAGE:
English
?logoff
>>>Unrecognizable Command
?logoff

21sep94 07:43:21 User214323 Session D226.4

\$2.83	0.008 Hrs	File398
\$24.70	19 Type(s)	in Format 3
\$24.70	19 Types	
\$0.00	View Fee	
\$27.53	Estimated cost	File398
\$2.30	0.024 Hrs	File5
\$4.50	5 Type(s)	in Format 3
\$12.60	14 Type(s)	in Format 7
\$17.10	19 Types	
\$0.00	View Fee	
\$19.40	Estimated cost	File5
\$2.59	0.016 Hrs	File434
\$0.00	1 Type(s)	in Format 55
\$0.00	1 Types	
\$0.00	View Fee	
\$2.59	Estimated cost	File434
\$9.90	0.275 Hrs	File155
\$0.12	1 Type(s)	in Format 3
\$1.68	14 Type(s)	in Format 7
\$1.80	15 Types	
\$0.00	View Fee	
\$11.70	Estimated cost	File155
\$0.40	0.011 Hrs	File159
\$0.24	2 Type(s)	in Format 7
\$0.24	2 Types	
\$0.00	View Fee	
\$0.64	Estimated cost	File159

\$12.38 0.086 Hrs File399
 \$24.20 22 Type(s) in Format 3
 \$2.20 2 Type(s) in Format 7
 \$26.40 24 Types
 \$0.00 View Fee
 \$38.78 Estimated cost File399
 \$1.02 0.010 Hrs File440
 \$0.00 View Fee
 \$1.02 Estimated cost File440
 \$2.70 0.025 Hrs File73
 \$15.00 15 Type(s) in Format 7
 \$15.00 15 Types
 \$0.00 View Fee
 \$17.70 Estimated cost File73
 \$0.66 0.010 Hrs File144
 \$0.00 View Fee
 \$0.66 Estimated cost File144
 \$0.90 0.010 Hrs File6
 \$1.50 2 Type(s) in Format 3
 \$1.50 2 Types
 \$0.00 View Fee
 \$2.40 Estimated cost File6
 \$0.81 0.005 Hrs File358
 \$0.00 View Fee
 \$0.81 Estimated cost File358
 OneSearch, 11 files, 0.483 Hrs FileOS
 \$5.51 TYMNET
 \$128.74 Estimated cost this search
 \$135.59 Estimated total session cost 0.613 Hrs.
 Logoff: level 37.09.15 D 07:43:21

TYMNET: call cleared by request

please log in:

APS
9/21/84
SKL

s autocrine motility factor or autotaxin or ATX
110 AUTOCRINE
1984 MOTILITY
187086 FACTOR
1 AUTOCRINE MOTILITY FACTOR
(AUTOCRINE(W)MOTILITY(W)FACTOR)
0 AUTOTAXIN
49 ATX
L1 50 AUTOCRINE MOTILITY FACTOR OR AUTOTAXIN OR ATX

=> asaabbbod?

L2 15817 ANTIBOD?

=> s 11 and 12

L3 2 L1 AND L2

=> # 13 1-2 bib, ab

US PAT NO: 5,248,596 [IMAGE AVAILABLE] L3: 1 of 2
DATE ISSUED: Sep. 28, 1993
TITLE: Method of detecting proteolytically modified antithrombin
INVENTOR: Pamela C. Esmon, Richmond, CA
Emma Yee, Albany, CA
Robert E. Jordan, Malvern, PA
Richard M. Nelson, La Jolla, CA
ASSIGNEE: Miles Inc., Berkeley, CA (U.S. corp.)
APPL-NO: 07/844,354
DATE FILED: Mar. 2, 1992
ART-UNIT: 182
PRIM-EXMR: Esther L. Kepplinger
ASST-EXMR: Carol E. Bidwell
LEGAL-REP: Elizabeth F. Enayati

US PAT NO: 5,248,596 [IMAGE AVAILABLE] L3: 1 of 2

ABSTRACT:

Indirect method of detecting elastase-modified or cleaved, antithrombin (**ATx**) in the presence of intact antithrombin (AT-III). The inventive method includes a modified ELISA using a detergent to alter the intact AT-III. Cleaved AT-III is generated in human plasma, then an ELISA is performed in the presence of a detergent.

US PAT NO: 5,245,022 [IMAGE AVAILABLE] L3: 2 of 2
DATE ISSUED: Sep. 14, 1993
TITLE: Exonuclease resistant terminally substituted
oligonucleotides
INVENTOR: Alexander L. Weis, Berwyn, PA
Fred T. Oakes, Rochester, NY
Frederick H. Hausheer, San Antonio, TX
Paul F. Cavanaugh, Jr., West Chester, PA
Patricia S. Moskwa, Phoenixville, PA

ASSIGNEE: Sterling Drug, Inc., New York, NY (U.S. corp.)
APPL-NO: 07/562,180
DATE FILED: Aug. 3, 1990
ART-UNIT: 183
PRIM-EXMR: John W. Rollins
ASST-EXMR: L. Eric Crane
LEGAL-REP: Irving Newman

US PAT NO: 5,245,022 [IMAGE AVAILABLE]

L3: 2 of 2

ABSTRACT:

Compounds, compositions and methods for inhibiting gene expression are disclosed. The compounds comprise oligonucleotide sequences of from about 9 to about 200 bases having a diol at either or both termini. Preferred diols are polyalkyleneglycols, preferably polyethyleneglycols. Pharmaceutical compositions comprising the compounds and a physiologically acceptable carrier and methods of inhibiting gene expression in mammals comprising administering such compounds are also provided. Methods for inhibiting nuclease cleavage of compounds are also provided.

=> d his

(FILE 'USPAT' ENTERED AT 08:42:36 ON 21 SEP 94)

L1 50 S AUTOCRINE MOTILITY FACTOR OR AUTOTAXIN OR ATX
L2 15817 S ANTIBOD?
L3 2 S L1 AND L2

=> d 11 not 13

L4 48 L1 NOT L3

=> d 14 1-10 bib,ab

US PAT NO: 5,325,985 [IMAGE AVAILABLE] L4: 1 of 48
DATE ISSUED: Jul. 5, 1994
TITLE: Gasket with a self-supporting protrusion
INVENTOR: Charles S. Radtke, Little Ferry, NJ
ASSIGNEE: Precision Valve Corporation, Yonkers, NY (U.S. corp.)
APPL-NO: 08/140,966
DATE FILED: Oct. 25, 1993
ART-UNIT: 241
PRIM-EXMR: Gary E. Elkins
LEGAL-REP: Davis Hoxie Faithfull & Hapgood

US PAT NO: 5,325,985 [IMAGE AVAILABLE]

L4: 1 of 48

ABSTRACT:

A gasket for sealing a channel of a mounting cup to a container bead including at least one protrusion at an end of the gasket for retaining the mounting cup in position on a container bead prior to clinching. Six such protrusions can be equidistantly positioned about the gasket. A gasketed mounting cup including such protrusions is also disclosed. A method for manufacturing such a gasketed mounting cup is disclosed as well.

US PAT NO: 5,321,089 [IMAGE AVAILABLE] L4: 2 of 48
DATE ISSUED: Jun. 14, 1994
TITLE: Golf ball cover
INVENTOR: Lauro C. Cadorniga, Piedmont, SC
Frank M. Simonutti, Anderson, SC
ASSIGNEE: Dunlop Slazenger Corporation, Greenville, SC (U.S. corp.)
APPL-NO: 08/039,902
DATE FILED: Mar. 30, 1993
ART-UNIT: 152

PRIM-EXMR: Carman J. Securo, Jr.
LEGAL-REP: Lorusso & Loud

US PAT NO: 5,321,089 [IMAGE AVAILABLE]

L4: 2 of 48

ABSTRACT:

A composition useful to produce a golf ball cover of a blend of ethylene-methyl acrylate and an ionomer resin and a compatibilizer. The composition produces a cover material having hardness and feel comparable to Balata, but having improved resilience, durability and cut resistance over Balata.

US PAT NO: 5,319,453 [IMAGE AVAILABLE]

L4: 3 of 48

DATE ISSUED: Jun. 7, 1994

TITLE: Method and apparatus for video signal encoding, decoding
and monitoring

INVENTOR: Robert C. Copriviza, Tarzana, CA
Arnold M. Dubin, Calabasas, CA
Edward B. Ackerman, Encino, CA
Jackson B. Wood, Tarzana, CA
Jeffrey S. Eakins, Claremont, CA
David D. Harmon, Torrance, CA

ASSIGNEE: Airtrax, Calabasas, CA (U.S. corp.)

APPL-NO: 07/370,399

DATE FILED: Jun. 22, 1989

ART-UNIT: 262

PRIM-EXMR: James J. Groody

ASST-EXMR: David E. Harvey

LEGAL-REP: Poms, Smith, Lande & Rose

US PAT NO: 5,319,453 [IMAGE AVAILABLE]

L4: 3 of 48

ABSTRACT:

Unique digital codes are encoded on a video signal, the codes are retrieved at receivers and precise information concerning the time of occurrence, length, nature and quality of a monitored broadcast at a frame by frame level, is generated. The codes are inserted on scan lines of the video, and vary either on a field-to-field or frame-to-frame basis. The code has a repeating first part having a unique program material identifier indicating the time, date and place of encoding, and has a second portion that varies in a predetermined non-repeating sequence which varies along the entire length of the tape, thereby uniquely identifying each frame of the video program material. Also encoded upon successive frames is a cyclic counter code with a count corresponding to the sequence of the identifier data on successive frames. When the video signal is processed by a receiver, the first portion identifier data from the various frames is mapped into selected memory locations in accordance with the count of the frame as determined by the second portion. Odd and even fields are encoded with complementary bit sequences to assist in processing the encoded data. Whenever the frame sequence is interrupted a data packet is generated representative of the condition encountered. The data packets are accumulated in log files in a memory in the receiver. The log files are transmitted to a data center, as is a copy of the encoded tape. Reports concerning the broadcast are generated.

US PAT NO: 5,317,391 [IMAGE AVAILABLE]

L4: 4 of 48

DATE ISSUED: May 31, 1994

TITLE: Method and apparatus for providing message information to
subscribers in a cable television system

INVENTOR: Robert O. Banker, Cumming, GA
Kinney C. Bacon, Lawrenceville, GA
Julius B. Bagley, Marietta, GA

ASSIGNEE: Scientific-Atlanta, Inc., Norcross, GA (U.S. corp.)

APPL-NO: 07/799,987
DATE FILED: Nov. 29, 1991
ART-UNIT: 261
PRIM-EXMR: Reinhard J. Eisenzopf
ASST-EXMR: Nguyen Vo
LEGAL-REP: Banner, Birch, McKie & Beckett

US PAT NO: 5,317,391 [IMAGE AVAILABLE]

L4: 4 of 48

ABSTRACT:

A subscriber terminal apparatus for a television in an in-band subscription television system is provided. The subscriber terminal includes a receiver for receiving a television signal including video, audio, and data information. A selector selects a channel of the television signal. A memory stores a plurality of barker screens providing messages regarding one or more channels of the television signal. An on-screen display control circuit controls the display of the barker screens on the television. A processor retrieving a barker screen from the memory supplies the retrieved screen to the on-screen display control circuit if the barker screen provides a message regarding a selected channel. Barker screen information may also be obtained from a dedicated data channel or a six megahertz video barker channel.

US PAT NO: 5,309,514 [IMAGE AVAILABLE]

L4: 5 of 48

DATE ISSUED: May 3, 1994

TITLE: Pulse generator including a memory for storing pulses for modulation on a carrier of a television signal

INVENTOR: Marshall B. Johnson, Norcross, GA
Lamar E. West, Jr., Maysville, GA

ASSIGNEE: Scientific-Atlanta, Inc., Atlanta, GA (U.S. corp.)

APPL-NO: 07/891,053

DATE FILED: Jun. 1, 1992

ART-UNIT: 222

PRIM-EXMR: Tod R. Swann

LEGAL-REP: Banner, Birch, McKie Beckett

US PAT NO: 5,309,514 [IMAGE AVAILABLE]

L4: 5 of 48

ABSTRACT:

A pulse generator for generating pulses for modulation onto a carrier or subcarrier of a composite television signal is provided. The pulse generator includes a memory such as an EPROM for storing one or more waveshapes. Each stored waveshape is defined by a sequence of addressable values representing the amplitude of the waveshape as a function of time. A selecting circuit such as a microprocessor selects one of the waveshapes in the memory. A counting circuit responsive to a clock signal controls the address lines of the memory to read the amplitude values corresponding to the selected waveshape from the memory. The amplitude values are supplied to a digital to analog converter to convert the amplitude values to an analog pulse. The pulse may then be filtered to remove clock noise. The resultant signal is supplied to an amplitude modulator for modulating the signal onto a carrier or subcarrier of a composite television signal.

US PAT NO: 5,301,028 [IMAGE AVAILABLE]

L4: 6 of 48

DATE ISSUED: Apr. 5, 1994

TITLE: Method and apparatus for displaying channel identification information

INVENTOR: Robert O. Banker, Cumming, GA
Kinney C. Bacon, Lawrenceville, GA
Julius B. Bagley, Marietta, GA

ASSIGNEE: Scientific-Atlanta, Inc., Norcross, GA (U.S. corp.)

APPL-NO: 07/800,002

DATE FILED: Nov. 29, 1991

ART-UNIT: 262
PRIM-EXMR: Mark R. Powell
ASST-EXMR: Jeffrey S. Murrell
LEGAL-REP: Frederick W. Powers, III

US PAT NO: 5,301,028 [IMAGE AVAILABLE]

L4: 6 of 48

ABSTRACT:

A subscriber terminal includes a receiver for receiving a television signal including video, audio, and data information. A channel of the television signal may be selected for display on the television. A memory stores channel identification information such as channel identifiers. The channel identifiers include at least one display character. A processor establishes a relationship between channel identification information and channel numbers associated with channels of the television signal. An on-screen display control circuit controls the display of character information on the television and has the capability of overlaying the channel number and the channel identification information on the video portion of a selected channel displayed on the television for a predetermined period of time. Program identification information such as program titles may also be displayed. Alternatively or additionally, the channel identification information may be displayed on a display such as an LED display of a subscriber terminal.

US PAT NO: 5,291,459 [IMAGE AVAILABLE]

L4: 7 of 48

DATE ISSUED: Mar. 1, 1994

TITLE: Signal processor having multiple distributed data buffers

INVENTOR: Victor A. Andersen, North Dartmouth, MA

ASSIGNEE: The United States of America as represented by the
Secretary of the Navy, Washington, DC (U.S. govt.)

APPL-NO: 08/059,770

DATE FILED: May 7, 1993

ART-UNIT: 221

PRIM-EXMR: Daniel T. Pihulic

LEGAL-REP: Michael J. McGowan, Prithvi C. Lall, Michael F. Oglo

US PAT NO: 5,291,459 [IMAGE AVAILABLE]

L4: 7 of 48

ABSTRACT:

A hydrophone analog signal data acquisition, A/D conversion and data transmission system includes a first-stage signal processing subsystem which provides digital representations of the hydrophone analog signal, which in turn are signal processed for transmission in the form of data packets by a second stage signal processing subsystem (40). Subsystem 40 includes a plurality of Data Multiplexer/FIFO units (48), including corresponding selectively acting data unit accumulators, each accumulator having a plurality of inputs coupled to output channels of the first-stage signal processing subsystem for receiving digital representations of hydrophone analog signals. Each data unit accumulator includes a first buffer (48-2) for storing information that includes a digital representation of the analog hydrophone signal, an identification of a hydrophone that generated the acoustic information, and a time that the acoustic information is received from the hydrophone. Each data unit accumulator further includes an input interface that is operable during the first period for receiving a alert signal with a hydrophone analog signal, indicating that the associated source has data available. The input interface compares a current state of the alert signal to a previous state for detecting an occurrence of the assertion of the alert signal. The data unit accumulator also receives and stores a unit of data from a data source having an asserted alert signal, and is responsive to the storage of the unit of data therein, during the first period, to receive and store, during the second period, other information associated with the unit of data stored during the first period.

US PAT NO: 5,286,541 [IMAGE AVAILABLE] L4: 8 of 48
DATE ISSUED: Feb. 15, 1994
TITLE: Coated abrasive having combination backing member
INVENTOR: Dhiraj H. Darjee, Ballston Lake, NY
Richard W. Kalita, Ballston Lake, NY
Gregg M. Bosak, Hoosick Falls, NY
Eugene Zador, Ballston Lake, NY
William F. McCutcheon, Mission, TX
ASSIGNEE: Norton Company, Worcester, MA (U.S. corp.)
APPL-NO: 07/943,077
DATE FILED: Sep. 10, 1992
ART-UNIT: 158
PRIM-EXMR: Ellis P. Robinson
ASST-EXMR: Nasser Ahmad
LEGAL-REP: David Bennett

US PAT NO: 5,286,541 [IMAGE AVAILABLE] L4: 8 of 48

ABSTRACT:

Coated abrasive material having a combination backing member. The combination backing member has as a bottom member a conventional backing member substrate used in the manufacture of coated abrasive material such as cylinder paper coated with a polymeric layer. The polymer layer provides a relatively smooth surface for application of the maker coat during manufacture of the coated abrasive material. The coated abrasive member can be used in diverse applications such as the fine finishing of particle board and offhand grinding of automobile body seams.

US PAT NO: 5,281,651 [IMAGE AVAILABLE] L4: 9 of 48
DATE ISSUED: Jan. 25, 1994
TITLE: Compatibilization of dissimilar elastomer blends using
ethylene/acrylate/acrylic acid terpolymers
INVENTOR: Palanisamy Arjunan, Dayton, NJ
Roma B. Kuszniir, Flushing, NY
ASSIGNEE: Exxon Chemical Patents Inc., Linden, NJ (U.S. corp.)
APPL-NO: 07/827,772
DATE FILED: Jan. 29, 1992
ART-UNIT: 152
PRIM-EXMR: Carman J. Seccuro, Jr.
LEGAL-REP: Catherine L. Bell

US PAT NO: 5,281,651 [IMAGE AVAILABLE] L4: 9 of 48

ABSTRACT:

This invention relates to a compatibilized rubber composition and a process for compatibilizing dissimilar rubber blends comprising blending an ethylene/acrylate/acrylic acid terpolymer with 2 or more different rubbers, selected from the group including, but not limited to, EPR, EPDM, CR, NBR, SBR and NR.

US PAT NO: 5,255,086 [IMAGE AVAILABLE] L4: 10 of 48
DATE ISSUED: Oct. 19, 1993
TITLE: Method and apparatus for RF data transfer in a CATV system
INVENTOR: Jay C. McMullan, Jr., Doraville, GA
David J. Naddor, Doraville, GA
Robert J. Beyers, II, Snellville, GA
ASSIGNEE: Scientific-Atlanta, Inc., Norcross, GA (U.S. corp.)
APPL-NO: 07/562,675
DATE FILED: Aug. 3, 1990
ART-UNIT: 261
PRIM-EXMR: Reinhard J. Eisenzopf
ASST-EXMR: Chi H. Pham
LEGAL-REP: William A. Marvin, Frederick W. Powers, III

ABSTRACT:

A method of controlling the allocation of a population of remote units among a plurality of groups of remote units is provided. Each remote unit has a digital identifier respectively associated therewith. A maximum and a minimum average number of remote units per group is fixed. The remote units are assigned to the groups of remote units in accordance with the respective digital identifiers. The average number of remote units per group is then determined as remote units are assigned thereto. Next, the average number of remote units per group is compared to the fixed maximum number of remote units per group. The above steps are repeated while the average number of remote units per group is less than or equal to the fixed maximum number of remote units per group. The number of groups is changed such that the average number of remote units per group is between the fixed maximum and minimum number of remote units per group if the average number of remote units per group exceeds the maximum number of remote units per group.

=> d 14 11-48 ti

US PAT NO:	5,251,324 [IMAGE AVAILABLE]	L4: 11 of 48
TITLE:	Method and apparatus for generating and collecting viewing statistics for remote terminals in a cable television system	
US PAT NO:	5,247,364 [IMAGE AVAILABLE]	L4: 12 of 48
TITLE:	Method and apparatus for tuning data channels in a subscription television system having in-band data transmissions	
US PAT NO:	5,236,636 [IMAGE AVAILABLE]	L4: 13 of 48
TITLE:	In-mold plasma treatment	
US PAT NO:	5,235,619 [IMAGE AVAILABLE]	L4: 14 of 48
TITLE:	Cable television radio frequency subscriber data transmission apparatus and RF return method	
US PAT NO:	5,228,034 [IMAGE AVAILABLE]	L4: 15 of 48
TITLE:	Ring communication network station	
US PAT NO:	5,227,180 [IMAGE AVAILABLE]	L4: 16 of 48
TITLE:	Apparatus for applying an electric field	
US PAT NO:	5,225,902 [IMAGE AVAILABLE]	L4: 17 of 48
TITLE:	Automatic frequency selection in a bi-directional cable television system	
US PAT NO:	5,175,766 [IMAGE AVAILABLE]	L4: 18 of 48
TITLE:	Signalling scheme for controlling data encryption device in an electronic fund transaction processing system	
US PAT NO:	5,168,714 [IMAGE AVAILABLE]	L4: 19 of 48
TITLE:	Assembly, especially for a beverage-vending machine, with a container for the storage, cooling and carbonating of water	
US PAT NO:	5,155,590 [IMAGE AVAILABLE]	L4: 20 of 48
TITLE:	System for data channel level control	
US PAT NO:	5,142,690 [IMAGE AVAILABLE]	L4: 21 of 48
TITLE:	Cable television radio frequency data processor	
US PAT NO:	5,132,315 [IMAGE AVAILABLE]	L4: 22 of 48
TITLE:	Therapeutic application of an anti-invasive compound	

US PAT NO:	5,058,160 [IMAGE AVAILABLE]	L4: 23 of 48
TITLE:	In-band controller	
US PAT NO:	5,045,816 [IMAGE AVAILABLE]	L4: 24 of 48
TITLE:	Binary phase shift key modulator with programmable level control	
US PAT NO:	5,014,315 [IMAGE AVAILABLE]	L4: 25 of 48
TITLE:	Digital telephone system	
US PAT NO:	5,012,510 [IMAGE AVAILABLE]	L4: 26 of 48
TITLE:	Dynamic callback technique	
US PAT NO:	5,003,384 [IMAGE AVAILABLE]	L4: 27 of 48
TITLE:	Set-top interface transactions in an impulse pay per view television system	
US PAT NO:	4,922,456 [IMAGE AVAILABLE]	L4: 28 of 48
TITLE:	Method of reducing wearout in a non-volatile memory with double buffer	
US PAT NO:	4,798,543 [IMAGE AVAILABLE]	L4: 29 of 48
TITLE:	Interactive training method and system	
US PAT NO:	4,773,378 [IMAGE AVAILABLE]	L4: 30 of 48
TITLE:	Fuel supply control method for internal combustion engines after starting in hot state	
US PAT NO:	4,724,340 [IMAGE AVAILABLE]	L4: 31 of 48
TITLE:	Output circuit in which induced switching noise is reduced by presetting pairs of output lines to opposite logic states	
US PAT NO:	4,692,549 [IMAGE AVAILABLE]	L4: 32 of 48
TITLE:	Carboalkoxylation of butadiene to form dialkyl adipate	
US PAT NO:	4,576,578 [IMAGE AVAILABLE]	L4: 33 of 48
TITLE:	Interactive training apparatus	
US PAT NO:	4,570,035 [IMAGE AVAILABLE]	L4: 34 of 48
TITLE:	Programmable key telephone system	
US PAT NO:	4,569,062 [IMAGE AVAILABLE]	L4: 35 of 48
TITLE:	Interface circuit for interfacing between asynchronous data in start/stop format and synchronous data	
US PAT NO:	4,561,088 [IMAGE AVAILABLE]	L4: 36 of 48
TITLE:	Communication system bypass architecture	
US PAT NO:	4,554,413 [IMAGE AVAILABLE]	L4: 37 of 48
TITLE:	Key telephone system	
US PAT NO:	4,535,084 [IMAGE AVAILABLE]	L4: 38 of 48
TITLE:	Certain 4-(2-hydroxyethylthiomethyl)pyridines and derivatives thereof having immunoregulatory activity	
US PAT NO:	4,500,535 [IMAGE AVAILABLE]	L4: 39 of 48
TITLE:	Method of regulating the immune response with pyridine derivatives	
US PAT NO:	4,371,871 [IMAGE AVAILABLE]	L4: 40 of 48
TITLE:	Alert message communication system	
US PAT NO:	4,371,696 [IMAGE AVAILABLE]	L4: 41 of 48
TITLE:	Certain pyridine methylthio acetaldehyde derivatives and non-cyclic and cyclic acetals thereof	

US PAT NO: 4,354,241 [IMAGE AVAILABLE] L4: 42 of 48
 TITLE: Programmable electronic real-time load controller
 providing for adaptation of load control in response to
 varying environmental conditions

US PAT NO: 4,246,263 [IMAGE AVAILABLE] L4: 43 of 48
 TITLE: Antiinflammatory and immunoregulatory pyrimidines, their
 method of use and pharmaceutical compositions

US PAT NO: 4,191,953 [IMAGE AVAILABLE] L4: 44 of 48
 TITLE: Intrusion sensor and aerial therefor

US PAT NO: 4,138,718 [IMAGE AVAILABLE] L4: 45 of 48
 TITLE: Numerical control system with downloading capability

US PAT NO: 3,939,148 [IMAGE AVAILABLE] L4: 46 of 48
 TITLE: Process for preparing 1,3,5,7-tetranitro-1,3,5,7-
 tetraazacyclooctane

US PAT NO: 3,766,370 [IMAGE AVAILABLE] L4: 47 of 48
 TITLE: ELEMENTARY FLOATING POINT CORDIC FUNCTION PROCESSOR AND
 SHIFTER

US PAT NO: 3,632,883 [IMAGE AVAILABLE] L4: 48 of 48
 TITLE: TELECOMMUNICATION EXCHANGE WITH TIME DIVISION MULTIPLEX

=> d 14 43 bib,ab

US PAT NO: 4,246,263 [IMAGE AVAILABLE] L4: 43 of 48
 DATE ISSUED: Jan. 20, 1981
 TITLE: Antiinflammatory and immunoregulatory pyrimidines, their
 method of use and pharmaceutical compositions

INVENTOR: Joseph G. Lombardino, Groton, CT
 Charles A. Harbert, Groton, CT

ASSIGNEE: Pfizer Inc., New York, NY (U.S. corp.)

APPL-NO: 06/085,011

DATE FILED: Oct. 15, 1979

ART-UNIT: 121

PRIM-EXMR: Henry R. Jiles

ASST-EXMR: Robert T. Bond

LEGAL-REP: Connolly and Hutz

US PAT NO: 4,246,263 [IMAGE AVAILABLE] L4: 43 of 48

ABSTRACT:

A series of 4-(2-hydroxyethylthiomethyl)pyridines and related compounds, and their pharmaceutically acceptable acid addition salts, having antiinflammatory and immunoregulatory activity are disclosed. Preferred compounds include 4-(2-hydroxyethylthiomethyl)pyridine itself, as well as 4-(2-hydroxyphenylthiomethyl)pyridine, 4-(2-hydroxy-1-propylthiomethyl)pyridine, 4-(3-hydroxy-2-butylthiomethyl)pyridine, 4-(2-hydroxyethylthiomethyl)pyrimidine, the acetate esters corresponding to the above compounds, and 4-(2,3-dihydroxy-1-propylthiomethyl)pyridine.

=> log y

U.S. Patent & Trademark Office LOGOFF AT 08:49:31 ON 21 SEP 94

=> fil reg; s whvar | pldvky | ypafk | qaevs | peevtrpnyl | ydvpwneti |
vppfeniely | gggplwitatk | vnsmtvfvgygptfk | diehltsldffr |
teflsnyltnvddit/sqsp

FILE 'REGISTRY' ENTERED AT 14:12:16 ON 21 SEP 94

USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT

COPYRIGHT (C) 1994 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 17 SEP 94 HIGHEST RN 157748-89-5

DICTIONARY FILE UPDATES: 20 SEP 94 HIGHEST RN 157748-89-5

TSCA INFORMATION NOW CURRENT THROUGH MAY 1994

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

L1 35 WHVAR|PLDVYK|YPAFK|QAEVS|PEEVTRPNYL|YDVPWNETI|VPPFENIELY|G
GQPLWITATK|VNSMTVFPVGYGPTFK|DIEHLTSLDFFR|TEFLSNYLTNVDDIT/S
QSP

=> fil ca; s l1 and (autotox? or auto tox? or atx# or amf# or
autocrin?(1w)factor# or motility(1w)autocrin? and antibod?)/ab,bi

FILE 'CA' ENTERED AT 14:13:21 ON 21 SEP 94

USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT

COPYRIGHT (C) 1994 AMERICAN CHEMICAL SOCIETY (ACS)

FILE COVERS 1967 - 17 Sep 1994 (940917/ED) VOL 121 ISS 12

To help control your online searching costs, consider using the
HCA File when conducting SmartSELECT searches with large
numbers of terms.

THE BASIC INDEX NOW INCLUDES ABSTRACT TEXT. SEE NEWS FOR DETAILS.

27 L1
47 AUTOTOX?/AB
67 AUTOTOX?/BI
3856 AUTO/AB
172776 TOX?/AB
0 AUTO TOX?/AB
((AUTO(W)TOX?)/AB)
16925 AUTO/BI
283181 TOX?/BI
1 AUTO TOX?/BI
((AUTO(W)TOX?)/BI)
130 ATX#/AB
135 ATX#/BI
270 AMF#/AB
309 AMF#/BI
3082 AUTOCRIN?/AB
3260 AUTOCRIN?/BI
443319 FACTOR#/AB

524327 FACTOR#/BI
 616 AUTOCRIN?(1W)FACTOR#
 9802 MOTILITY/AB
 12013 MOTILITY/BI
 3082 AUTOCRIN?/AB
 3260 AUTOCRIN?/BI
 2 MOTILITY(1W)AUTOCRIN?
 157273 ANTIBOD?/AB
 173138 ANTIBOD?/BI
 L2 1 L1 AND (AUTOTOX? OR AUTO TOX? OR ATX# OR AMF# OR AUTOCRIN?
 (1W)FACTOR# OR MOTILITY(1W)AUTOCRIN? AND ANTIBOD?)/AB,BI

=> d .bevstr; sel hit 12 rn

L2 ANSWER 1 OF 1 CA COPYRIGHT 1994 ACS
 AN 119:218353 CA
 TI Autotaxin: motility stimulating protein useful in cancer diagnosis
 and therapy.
 IN Stracke, Mary; Liotta, Lance A.; Schiffmann, Elliott; Kratzsch,
 Henry
 PA United States Dept. of Health and Human Services, USA
 SO U. S. Pat. Appl., 61 pp. Avail. NTIS Order No. PAT-APPL-7-822 043.
 CODEN: XAXXAV
 PI US 822043 A0 930101
 AI US 92-822043 920117
 DT Patent
 LA English
 AB Autotaxin (ATX), an autocrine factor
 produced by A2058 human melanoma cells, is purified and
 characterized. ATX appears to be a glycosylated protein,
 has a pI of 7.7, and has a mol. wt. of 125 kDa on SDS-PAGE under
 reducing conditions. Purified ATX is active in the
 picomolar range. Peptide fragments of ATX were sequenced:
 no significant homologies to these peptides were found when
 searching GenBank, EMBL, SWISS-PROT, or GenPept protein databases.
 IT 147960-52-9, Autotaxin fragment (human) 147960-53-0
 , Autotaxin fragment (human) 147960-54-1, Autotaxin
 fragment (human) 147960-56-3, Autotaxin fragment (human)
 147960-57-4, Autotaxin fragment (human) 147960-58-5
 , Autotaxin fragment (human) 147977-73-9, Autotaxin
 fragment (human)
 (fragment of human autotaxin)

E1 THROUGH E7 ASSIGNED

=> fil reg; s e1-e7

FILE 'REGISTRY' ENTERED AT 14:14:00 ON 21 SEP 94
 USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
 COPYRIGHT (C) 1994 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 17 SEP 94 HIGHEST RN 157748-89-5
 DICTIONARY FILE UPDATES: 20 SEP 94 HIGHEST RN 157748-89-5

TSCA INFORMATION NOW CURRENT THROUGH MAY 1994

Please note that search-term pricing does apply when conducting SmartSELECT searches.

1 147960-52-9/BI
(147960-52-9/RN)
1 147960-53-0/BI
(147960-53-0/RN)
1 147960-54-1/BI
(147960-54-1/RN)
1 147960-56-3/BI
(147960-56-3/RN)
1 147960-57-4/BI
(147960-57-4/RN)
1 147960-58-5/BI
(147960-58-5/RN)
1 147977-73-9/BI
(147977-73-9/RN)

L3 7 (147960-52-9/BI OR 147960-53-0/BI OR 147960-54-1/BI OR 147960-56-3/BI OR 147960-57-4/BI OR 147960-58-5/BI OR 147977-73-9/BI)

=> d 1-7 .bevreg; fil caprev; s 11

L3 ANSWER 1 OF 7 REGISTRY COPYRIGHT 1994 ACS
RN 147977-73-9 REGISTRY
CN L-Lysine, N2-[N-[N-[N-[N-[N-[1-[N2-(N-glycylglycyl)-L-glutaminy]-L-prolyl]-L-leucyl]-L-tryptophyl]-L-isoleucyl]-L-threonyl]-L-alanyl]-L-threonyl]- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 11
MF C54 H86 N14 O15

L3 ANSWER 2 OF 7 REGISTRY COPYRIGHT 1994 ACS
RN 147960-58-5 REGISTRY
CN L-Arginine, L-threonyl-L-.alpha.-glutamyl-L-phenylalanyl-L-leucyl-L-seryl-L-asparaginy-L-tyrosyl-L-leucyl-L-threonyl-L-asparaginy-L-valyl-L-.alpha.-aspartyl-L-.alpha.-aspartyl-L-isoleucyl-L-threonyl-L-leucyl-L-valyl-L-prolylglycyl-L-threonyl-L-leucylglycyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 23
MF C113 H180 N28 O38

L3 ANSWER 3 OF 7 REGISTRY COPYRIGHT 1994 ACS
RN 147960-57-4 REGISTRY
CN L-Arginine, N2-[N-[N-[N-[N-[N-[N-[N-[N-(N-L-.alpha.-aspartyl-L-isoleucyl)-L-.alpha.-glutamyl]-L-histidyl]-L-leucyl]-L-threonyl]-L-

seryl]-L-leucyl]-L-.alpha.-aspartyl]-L-phenylalanyl]-L-phenylalanyl]-
(9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 12
MF C68 H101 N17 O21

L3 ANSWER 4 OF 7 REGISTRY COPYRIGHT 1994 ACS

RN 147960-56-3 REGISTRY

CN L-Lysine, L-valyl-L-asparaginyl-L-seryl-L-methionyl-L-glutaminyl-L-threonyl-L-valyl-L-phenylalanyl-L-valylglycyl-L-tyrosylglycyl-L-prolyl-L-threonyl-L-phenylalanyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 16
MF C82 H123 N19 O23 S

L3 ANSWER 5 OF 7 REGISTRY COPYRIGHT 1994 ACS

RN 147960-54-1 REGISTRY

CN L-Isoleucine, N-[N-[N-[N2-[N-[1-[N-(N-L-tyrosyl-L-.alpha.-aspartyl)-L-valyl]-L-prolyl]-L-tryptophyl]-L-asparaginyl]-L-.alpha.-glutamyl]-L-threonyl]- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 9
MF C53 H73 N11 O17

L3 ANSWER 6 OF 7 REGISTRY COPYRIGHT 1994 ACS

RN 147960-53-0 REGISTRY

CN L-Serine, N-[N-[N-(N-L-glutaminyl-L-alanyl)-L-.alpha.-glutamyl]-L-valyl]- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 5
MF C21 H36 N6 O10

L3 ANSWER 7 OF 7 REGISTRY COPYRIGHT 1994 ACS

RN 147960-52-9 REGISTRY

CN L-Lysine, N2-[N-[N-(1-L-tyrosyl-L-prolyl)-L-alanyl]-L-phenylalanyl]- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Autotaxin fragment (human)
SQL 5
MF C32 H44 N6 O7

FILE 'CAPREVIEWS' ENTERED AT 14:14:33 ON 21 SEP 94
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
COPYRIGHT (C) 1994 American Chemical Society (ACS)

FILE COVERS CURRENT RECORDS AND IS UPDATED DAILY
FILE LAST UPDATED: 20 SEP 1994 (940920/ED)

Some records may not contain a volume, issue, and page number, because they are from advance abstracts of articles accepted for publication in journals published by the ACS. These abstracts will appear first in Advance ACS Abstracts. Please enter HELP ACSJOURNALS to see a list of the advance ACS journal titles.

To help control your online searching costs, consider using the HCAPREVIEWS File when conducting SmartSELECT searches with large numbers of terms.

L4 0 L1

=> fil reg; s l1 not l3

FILE 'REGISTRY' ENTERED AT 14:14:49 ON 21 SEP 94
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
COPYRIGHT (C) 1994 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 17 SEP 94 HIGHEST RN 157748-89-5
DICTIONARY FILE UPDATES: 20 SEP 94 HIGHEST RN 157748-89-5

TSCA INFORMATION NOW CURRENT THROUGH MAY 1994

Please note that search-term pricing does apply when conducting SmartSELECT searches.

L5 28 L1 NOT L3

=> fil ca; s l5

FILE 'CA' ENTERED AT 14:15:01 ON 21 SEP 94
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
COPYRIGHT (C) 1994 AMERICAN CHEMICAL SOCIETY (ACS)

FILE COVERS 1967 - 17 Sep 1994 (940917/ED) VOL 121 ISS 12

To help control your online searching costs, consider using the HCA File when conducting SmartSELECT searches with large numbers of terms.

THE BASIC INDEX NOW INCLUDES ABSTRACT TEXT. SEE NEWS FOR DETAILS.

L6 26 L5

=> s l6 not l2

L7 26 L6 NOT L2

=> d 1-26 .bev; sel hit l7 1-26 rn

L7 ANSWER 1 OF 26 CA COPYRIGHT 1994 ACS
 AN 121:50253 CA
 TI Human intestinal VIP receptor: cloning and functional expression of
 two cDNA encoding proteins with different N-terminal domains
 SO Biochem. Biophys. Res. Commun. (1994), 200(2), 769-76
 CODEN: BBRCA9; ISSN: 0006-291X
 AU Couvineau, Alain; Rouyer-Fessard, Christiane; Darmoul, Dalila;
 Maoret, Jean Jose; Carrero, Isabel; Ogier-Denis, Eric; Laburthe,
 Marc
 PY 1994

L7 ANSWER 2 OF 26 CA COPYRIGHT 1994 ACS
 AN 121:28178 CA
 TI A transforming fragment within the direct repeat region of human
 herpesvirus type 6 that transactivates HIV-1
 SO Oncogene (1994), 9(4), 1167-75
 CODEN: ONCNES; ISSN: 0950-9232
 AU Thompson, Jerry; Choudhury, Sukhendra; Kashanchi, Fatah; Doniger,
 Jay; Berneman, Zwi; Frenkel, Niza; Rosenthal, Leonard J.
 PY 1994

L7 ANSWER 3 OF 26 CA COPYRIGHT 1994 ACS
 AN 120:237313 CA
 TI Sequencing and functional analysis of a 32,560 bp segment on the
 left arm of yeast chromosome II. Identification of 26 open reading
 frames, including the KIP1 and SEC17 genes
 SO Yeast (1993), 9(12), 1355-71
 CODEN: YESTE3; ISSN: 0749-503X
 AU Scherens, Bart; El Bakkoury, Mohamed; Vierendeels, Fabienne; Dubois,
 Evelyne; Messenguy, Francine
 PY 1993

L7 ANSWER 4 OF 26 CA COPYRIGHT 1994 ACS
 AN 120:209523 CA
 TI Molecular cloning and expression of cDNA for vasoactive intestinal
 polypeptide (VIP) receptor of rat
 SO Jpn. Kokai Tokkyo Koho, 14 pp.
 CODEN: JKXXAF
 IN Osada, Juichi; Ishihara, Takeshi
 AI JP 92-26607 920213
 PI JP 05255394 A2 931005 Heisei
 PY 1993

L7 ANSWER 5 OF 26 CA COPYRIGHT 1994 ACS
 AN 120:97538 CA
 TI HSP90 homolog from Madagascar periwinkle (*Catharanthus roseus*): cDNA
 sequence, regulation of protein expression and location in the
 endoplasmic reticulum
 SO Plant Mol. Biol. (1993), 23(3), 583-94
 CODEN: PMBIDB; ISSN: 0167-4412
 AU Schroeder, Gudrun; Beck, Markus; Eichel, Johannes; Vetter, Hans
 Peter; Schroeder, Joachim
 PY 1993

L7 ANSWER 6 OF 26 CA COPYRIGHT 1994 ACS
 AN 119:263762 CA
 TI The partial 3'-conserved segment duplications in the integrons In6
 from pSa and In7 from pDGO100 have a common origin
 SO Plasmid (1993), 30(1), 39-50
 CODEN: PLSMDX; ISSN: 0147-619X
 AU Stokes, H. W.; Tomaras, C.; Parsons, Yvonne; Hall, Ruth M.
 PY 1993

L7 ANSWER 7 OF 26 CA COPYRIGHT 1994 ACS
 AN 119:247982 CA
 TI Humanized monoclonal antibodies against human interleukin-4
 SO PCT Int. Appl., 115 pp.
 CODEN: PIXXD2
 IN Abrams, John S.; Dalie, Barbara; Le, Hung V.; Miller, Kenneth;
 Murgolo, Nicholas J.; Nguyen Hanh; Pearce, Michael; Tindall,
 Stephen; Zavodny, Paul J.
 AI WO 93-US1301 930218
 PI WO 9317106 A1 930902
 PY 1993

L7 ANSWER 8 OF 26 CA COPYRIGHT 1994 ACS
 AN 119:152250 CA
 TI Cloning and functional expression of a human neuroendocrine
 vasoactive intestinal peptide receptor
 SO Biochem. Biophys. Res. Commun. (1993), 193(2), 546-53
 CODEN: BBRCA9; ISSN: 0006-291X
 AU Sreedharan, Sunil P.; Patel, Derek R.; Huang, Jin Xing; Goetzl,
 Edward J.
 PY 1993

L7 ANSWER 9 OF 26 CA COPYRIGHT 1994 ACS
 AN 119:111597 CA
 TI A pathogen-induced gene of barley encodes a HSP90 homolog showing
 striking similarity to vertebrate forms resident in the endoplasmic
 reticulum
 SO Plant Mol. Biol. (1993), 21(6), 1097-108
 CODEN: PMBIDB; ISSN: 0167-4412
 AU Walther-Larsen, Haidee; Brandt, Jakob; Collinge, David B.;
 Thordal-Christensen, Hans
 PY 1993

L7 ANSWER 10 OF 26 CA COPYRIGHT 1994 ACS
 AN 119:43542 CA
 TI The primary structure of a protein containing a putative iron-sulfur
 [6Fe-6S] prismane cluster from Desulfovibrio vulgaris
 (Hildenborough)
 SO Eur. J. Biochem. (1992), 208(2), 435-42
 CODEN: EJBCAI; ISSN: 0014-2956
 AU Stokkermans, Jack P. W. G.; Pierik, Antonio J.; Wolbert, Ronnie B.
 G.; Hagen, Wilfred R.; Van Dongen, Walter M. A. M.; Veeger, Cees
 PY 1992

L7 ANSWER 11 OF 26 CA COPYRIGHT 1994 ACS
AN 119:22989 CA
TI The primary structure of a protein containing a putative iron-sulfur
[6Fe-6S] prismane cluster from *Desulfovibrio desulfuricans* (ATCC
27774)
SO *Biochim. Biophys. Acta* (1992), 1132(1), 83-7
CODEN: BBACAQ; ISSN: 0006-3002
AU Stokkermans, Jack P. W. G.; Van den Berg, Willy A. M.; Van Dongen,
Walter M. A. M.; Veeger, Cees
PY 1992

L7 ANSWER 12 OF 26 CA COPYRIGHT 1994 ACS
AN 118:249988 CA
TI Cytochrome P-450terp. Isolation and purification of the protein and
cloning and sequencing of its operon
SO *J. Biol. Chem.* (1992), 267(20), 14193-203
CODEN: JBCHA3; ISSN: 0021-9258
AU Peterson, Julian A.; Lu, Jui Yun; Geisselsoder, Janet;
Graham-Lorence, Sandra; Carmona, Cynthia; Witney, Frank; Lorence,
Matthew C.
PY 1992

L7 ANSWER 13 OF 26 CA COPYRIGHT 1994 ACS
AN 118:226995 CA
TI Complete nucleotide sequence of the *Actinomyces viscosus* T14V
sialidase gene: Presence of a conserved repeating sequence among
strains of *Actinomyces* spp
SO *Infect. Immun.* (1993), 61(1), 109-16
CODEN: INFIBR; ISSN: 0019-9567
AU Yeung, Maria K.
PY 1993

L7 ANSWER 14 OF 26 CA COPYRIGHT 1994 ACS
AN 118:226114 CA
TI Functional expression and tissue distribution of a novel receptor
for vasoactive intestinal polypeptide
SO *Neuron* (1992), 8(4), 811-19
CODEN: NERNET; ISSN: 0896-6273
AU Ishihara, Takeshi; Shigemoto, Ryuichi; Mori, Kensaku; Takahashi,
Kenji; Nagata, Shigekazu
PY 1992

L7 ANSWER 15 OF 26 CA COPYRIGHT 1994 ACS
AN 118:95239 CA
TI Six nodulation genes of nod box locus 4 in *Rhizobium meliloti* are
involved in nodulation signal production: nodM codes for
D-glucosamine synthetase
SO *Mol. Gen. Genet.* (1991), 228(1-2), 113-24
CODEN: MGGEAE; ISSN: 0026-8925
AU Baev, Nedelcho; Endre, Gabriella; Petrovics, Gyorgy; Banfalvi,
Zsafia; Kondorosi, Adam
PY 1991

L7 ANSWER 16 OF 26 CA COPYRIGHT 1994 ACS
 AN 118:2872 CA
 TI Cloning, sequencing and expression of the sialidase gene from
 Actinomyces viscosus DSM 43798
 SO Biol. Chem. Hoppe-Seyler (1991), 372(12), 1065-72
 CODEN: BCHSEI; ISSN: 0177-3593
 AU Henningsen, Michaela; Roggentin, Peter; Schauer, Roland
 PY 1991

L7 ANSWER 17 OF 26 CA COPYRIGHT 1994 ACS
 AN 117:68366 CA
 TI Chimeric and complementarity-determining region-grafted
 anti-carcinoembryonic antigen antibodies and their production
 SO PCT Int. Appl., 70 pp.
 CODEN: PIXXD2
 IN Adair, John Robert; Bodmer, Mark William; Mountain, Andrew; Owens,
 Raymond John
 AI WO 91-GB1108 910705
 PI WO 9201059 A1 920123
 PY 1992

L7 ANSWER 18 OF 26 CA COPYRIGHT 1994 ACS
 AN 116:1417 CA
 TI The regions of sequence variation in caulimovirus gene VI
 SO Virology (1991), 182(2), 830-4
 CODEN: VIRLAX; ISSN: 0042-6822
 AU Sanger, Margaret; Daubert, Steve; Goodman, Robert M.
 PY 1991

L7 ANSWER 19 OF 26 CA COPYRIGHT 1994 ACS
 AN 114:57704 CA
 TI The primary structure of DNA binding protein II from the extreme
 thermophilic bacterium Thermus thermophilus
 SO FEBS Lett. (1990), 273(1-2), 59-62
 CODEN: FEBLAL; ISSN: 0014-5793
 AU Zierer, Rainer; Choli, Dora
 PY 1990

L7 ANSWER 20 OF 26 CA COPYRIGHT 1994 ACS
 AN 113:93447 CA
 TI The complete cDNA and polypeptide sequences of human erythroid
 .alpha.-spectrin
 SO J. Biol. Chem. (1990), 265(8), 4434-43
 CODEN: JBCHA3; ISSN: 0021-9258
 AU Sahr, Kenneth E.; Laurila, Pekka; Kotula, Leszek; Scarpa, Alphonse
 L.; Coupal, Elaine; Leto, Thomas L.; Linnenbach, Alban J.;
 Winkelmann, John C.; Speicher, David W.; et al.
 PY 1990

L7 ANSWER 21 OF 26 CA COPYRIGHT 1994 ACS
 AN 113:92067 CA
 TI Structural and functional analysis of the mini-circle, a

transposable element of *Streptomyces coelicolor* A3(2)
SO Mol. Microbiol. (1989), 3(10), 1307-18
CODEN: MOMIEE; ISSN: 0950-382X
AU Henderson, D. J.; Lydiate, D. J.; Hopwood, D. A.
PY 1989

L7 ANSWER 22 OF 26 CA COPYRIGHT 1994 ACS
AN 113:18653 CA
TI Characterization of the *Schizosaccharomyces pombe* ral2 gene
implicated in activation of the ras1 gene product
SO Mol. Cell. Biol. (1989), 9(12), 5617-22
CODEN: MCEBD4; ISSN: 0270-7306
AU Fukui, Yasuhisa; Miyake, Sanae; Satoh, Misako; Yamamoto, Masayuki
PY 1989

L7 ANSWER 23 OF 26 CA COPYRIGHT 1994 ACS
AN 112:192838 CA
TI Rat P45017.alpha. from testis: characterization of a full-length
cDNA encoding a unique steroid hydroxylase capable of catalyzing
both .DELTA.4- and .DELTA.5-steroid-17,20-lyase reactions
SO Mol. Endocrinol. (1989), 3(6), 968-75
CODEN: MOENEN; ISSN: 0888-8809
AU Fevold, H. Richard; Lorence, Matthew C.; McCarthy, John L.; Trant,
John M.; Kagimoto, Masaaki; Waterman, Michael R.; Mason, J. Ian
PY 1989

L7 ANSWER 24 OF 26 CA COPYRIGHT 1994 ACS
AN 111:72207 CA
TI Rat testis P-45017.alpha. DNA: the deduced amino acid sequence,
expression and secondary structural configuration
SO Biochem. Biophys. Res. Commun. (1988), 157(2), 705-12
CODEN: BBRCA9; ISSN: 0006-291X
AU Namiki, Mikio; Kitamura, Masaya; Buczko, Ellen; Dufau, Maria L.
PY 1988

L7 ANSWER 25 OF 26 CA COPYRIGHT 1994 ACS
AN 111:51701 CA
TI Pullulanase gene of *Klebsiella* and its cloning and sequencing
SO Jpn. Kokai Tokkyo Koho, 12 pp.
CODEN: JKXXAF
IN Muroka, Yoshikatsu; Takizawa, Noboru; Katsuragi, Nobuhiro
AI JP 87-78355 870331
PI JP 63245676 A2 881012 Showa
PY 1988

L7 ANSWER 26 OF 26 CA COPYRIGHT 1994 ACS
AN 107:110141 CA
TI Entire nucleotide sequence of the pullulanase gene of *Klebsiella*
aerogenes W70
SO J. Bacteriol. (1987), 169(5), 2301-6
CODEN: JOBAA; ISSN: 0021-9193
AU Katsuragi, Nobuhiro; Takizawa, Noboru; Murooka, Yoshikatsu
PY 1987

E8 THROUGH E33 ASSIGNED

=> fil reg; s e8-e33

FILE 'REGISTRY' ENTERED AT 14:16:02 ON 21 SEP 94

USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT

COPYRIGHT (C) 1994 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 17 SEP 94 HIGHEST RN 157748-89-5

DICTIONARY FILE UPDATES: 20 SEP 94 HIGHEST RN 157748-89-5

TSCA INFORMATION NOW CURRENT THROUGH MAY 1994

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

1 110071-62-0/BI
 (110071-62-0/RN)
1 110071-63-1/BI
 (110071-63-1/RN)
1 121938-40-7/BI
 (121938-40-7/RN)
1 147477-78-9/BI
 (147477-78-9/RN)
1 121763-54-0/BI
 (121763-54-0/RN)
1 127831-57-6/BI
 (127831-57-6/RN)
1 128771-18-6/BI
 (128771-18-6/RN)
1 128909-10-4/BI
 (128909-10-4/RN)
1 131571-34-1/BI
 (131571-34-1/RN)
1 137800-99-8/BI
 (137800-99-8/RN)
1 142661-57-2/BI
 (142661-57-2/RN)
1 144130-50-7/BI
 (144130-50-7/RN)
1 144813-79-6/BI
 (144813-79-6/RN)
1 147446-26-2/BI
 (147446-26-2/RN)
1 147883-85-0/BI
 (147883-85-0/RN)
1 148349-17-1/BI
 (148349-17-1/RN)
1 148591-55-3/BI
 (148591-55-3/RN)
1 149408-27-5/BI

(149408-27-5/RN)
 1 150138-10-6/BI
 (150138-10-6/RN)
 1 151066-62-5/BI
 (151066-62-5/RN)
 1 151552-72-6/BI
 (151552-72-6/RN)
 1 152745-57-8/BI
 (152745-57-8/RN)
 1 154363-46-9/BI
 (154363-46-9/RN)
 1 155980-72-6/BI
 (155980-72-6/RN)
 1 156288-65-2/BI
 (156288-65-2/RN)
 1 156288-66-3/BI
 (156288-66-3/RN)
 L8 26 (110071-62-0/BI OR 110071-63-1/BI OR 121938-40-7/BI OR 147
 477-78-9/BI OR 121763-54-0/BI OR 127831-57-6/BI OR 128771-
 18-6/BI OR 128909-10-4/BI OR 131571-34-1/BI OR 137800-99-8
 /BI OR 142661-57-2/BI OR 144130-50-7/BI OR 144813-79-6/BI
 OR 147446-26-2/BI OR 147883-85-0/BI OR 148349-17-1/BI OR 1
 48591-55-3/BI OR 149408-27-5/BI OR 150138-10-6/BI OR 15106
 6-62-5/BI OR 151552-72-6/BI OR 152745-57-8/BI OR 154363-46
 -9/BI OR 155980-72-6/BI OR 156288-65-2/BI OR 156288-66-3/B
 I)

=> d 1-26 .bevreg; fil hom

L8 ANSWER 1 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 156288-66-3 REGISTRY
 CN Protein (human clone hIVR5 vasoactive intestinal peptide
 receptor-related precursor reduced) (9CI) (CA INDEX NAME)
 OTHER NAMES:
 CN Protein (human clone hIVR5 vasoactive intestinal peptide
 receptor-related precursor)
 SQL 495
 MF Unspecified
 CI MAN

L8 ANSWER 2 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 156288-65-2 REGISTRY
 CN Receptor, vasoactive intestinal polypeptide (human clone hIVR8
 precursor reduced) (9CI) (CA INDEX NAME)
 OTHER NAMES:
 CN Vasoactive intestinal peptide/PACAP receptor (human clone hIVR8
 precursor)
 SQL 460
 MF Unspecified
 CI MAN

L8 ANSWER 3 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 155980-72-6 REGISTRY

CN Protein (human herpes virus 6 strain U1102 clone pNF1022 331-amino acid reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Genbank x73675-derived protein

SQL 331

MF Unspecified

CI MAN

L8 ANSWER 4 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 154363-46-9 REGISTRY

CN Protein (Saccharomyces cerevisiae clone .alpha.1006.13 gene YBL0511 reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Protein (Saccharomyces cerevisiae chromosome II clone .alpha.1006.13 gene YBL0511)

SQL 418

MF Unspecified

CI MAN

L8 ANSWER 5 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 152745-57-8 REGISTRY

CN Protein HSP 90 (Catharanthus roseus heat shock precursor reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Protein HSP90 (Catharanthus roseus precursor)

SQL 817

MF Unspecified

CI MAN

L8 ANSWER 6 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 151552-72-6 REGISTRY

CN Protein (plasmid pDGO100 integron In7 341-amino acid reduced) (9CI) (CA INDEX NAME)

SQL 341

MF Unspecified

CI MAN

L8 ANSWER 7 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 151066-62-5 REGISTRY

CN INDEX NAME NOT YET ASSIGNED

SQL 140

MF Unspecified

CI MAN

L8 ANSWER 8 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 150138-10-6 REGISTRY

CN Receptor, vasoactive intestinal polypeptide (human clone LHT21 reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN VIP receptor (human HT29 cell clone LHT21)

SQL 457

MF Unspecified

CI MAN

L8 ANSWER 9 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 149408-27-5 REGISTRY

CN Protein HSP 90 (barley clone pBT6-1 precursor reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Heat-shock protein GRP 94 homolog (Hordeum vulgare clone pBT6-1 precursor)

SQL 809

MF Unspecified

CI MAN

L8 ANSWER 10 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 148591-55-3 REGISTRY

CN Protein (Desulfovibrio vulgaris clone pJSP9 prismane iron-sulfur core-containing reduced) (9CI) (CA INDEX NAME)

SQL 553

MF Unspecified

CI MAN

L8 ANSWER 11 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 148349-17-1 REGISTRY

CN Protein (Desulfovibrio desulfuricans clone pWBP81 prismane iron-sulfur core-containing reduced) (9CI) (CA INDEX NAME)

SQL 545

MF Unspecified

CI MAN

L8 ANSWER 12 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 147883-85-0 REGISTRY

CN Cytochrome P 450 (Pseudomonas clone pT3/pT1 isoform terp reduced) (9CI) (CA INDEX NAME)

SQL 428

MF Unspecified

CI MAN

L8 ANSWER 13 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 147477-78-9 REGISTRY

CN Receptor, vasoactive intestinal polypeptide (rat clone pV19 precursor reduced) (9CI) (CA INDEX NAME)

SQL 459

MF Unspecified

CI MAN

L8 ANSWER 14 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 147446-26-2 REGISTRY

CN Neuraminidase (Actinomyces viscosus clone pMY450-1 gene nanH precursor reduced) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Sialidase (Actinomyces viscosus strain T14V clone pMY450-1 precursor reduced)

SQL 901

MF Unspecified

CI MAN

L8 ANSWER 15 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 144813-79-6 REGISTRY

CN Neuraminidase (Actinomyces viscosus strain DSM 43798 reduced) (9CI) (CA INDEX NAME)

SQL 913

MF Unspecified

CI MAN

L8 ANSWER 16 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 144130-50-7 REGISTRY

CN Protein (Rhizobium meliloti clone pD85 gene nolF reduced) (9CI) (CA INDEX NAME)

SQL 367

MF Unspecified

CI MAN

L8 ANSWER 17 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 142661-57-2 REGISTRY

CN 1-145-Immunoglobulin G 1 (human-mouse clone pAL45 gH1-A5B7 .gamma.1-chain anti-antigen CEA reduced) (9CI) (CA INDEX NAME)

SQL 146

MF Unspecified

CI MAN

L8 ANSWER 18 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 137800-99-8 REGISTRY

CN Protein IBMP (figwort mosaic virus clone DxS reduced) (9CI) (CA INDEX NAME)

SQL 512

MF Unspecified

CI MAN

L8 ANSWER 19 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 131571-34-1 REGISTRY

CN Protein II (Thermus thermophilus strain HB8 DNA-binding) (9CI) (CA INDEX NAME)

SQL 95

MF C458 H781 N127 O129 S

CI MAN

L8 ANSWER 20 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 128909-10-4 REGISTRY

CN Spectrin (human clone .alpha.3/.alpha.37/.alpha.7 .alpha.-subunit precursor reduced) (9CI) (CA INDEX NAME)

SQL 2426

MF Unspecified

CI MAN

L8 ANSWER 21 OF 26 REGISTRY COPYRIGHT 1994 ACS

RN 128771-18-6 REGISTRY

CN Protein (Streptomyces coelicolor strain 3(2) minicircle element

122-amino acid) (9CI) (CA INDEX NAME)
 SQL 122
 MF Unspecified
 CI MAN

L8 ANSWER 22 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 127831-57-6 REGISTRY
 CN Protein (Schizosaccharomyces pombe gene ral2 reduced) (9CI) (CA INDEX NAME)
 SQL 611
 MF Unspecified
 CI MAN

L8 ANSWER 23 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 121938-40-7 REGISTRY
 CN Cytochrome P 450 (rat clone NA isoform 17.alpha. protein moiety reduced) (9CI) (CA INDEX NAME)
 SQL 507
 MF Unspecified
 CI MAN

L8 ANSWER 24 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 121763-54-0 REGISTRY
 CN 2-1077-Pullulanase (Klebsiella pneumoniae protein moiety reduced), N-(1-oxohexadecyl)- (9CI) (CA INDEX NAME)
 SQL 1076
 MF Unspecified
 CI MAN

L8 ANSWER 25 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 110071-63-1 REGISTRY
 CN Pullulanase (Klebsiella pneumoniae protein moiety reduced) (9CI) (CA INDEX NAME)
 SQL 1077
 MF Unspecified
 CI MAN

L8 ANSWER 26 OF 26 REGISTRY COPYRIGHT 1994 ACS
 RN 110071-62-0 REGISTRY
 CN Pullulanase (Klebsiella pneumoniae precursor protein moiety reduced) (9CI) (CA INDEX NAME)
 SQL 1096
 MF Unspecified
 CI MAN

FILE 'HOME' ENTERED AT 14:16:45 ON 21 SEP 94

Loring
249182
(Pt. 1-2)

Seq IDs 1-11 w/ Terms "autotoxin" or "ATX" or "AmF"
"autocrine" or motility"
searched in all AA databases

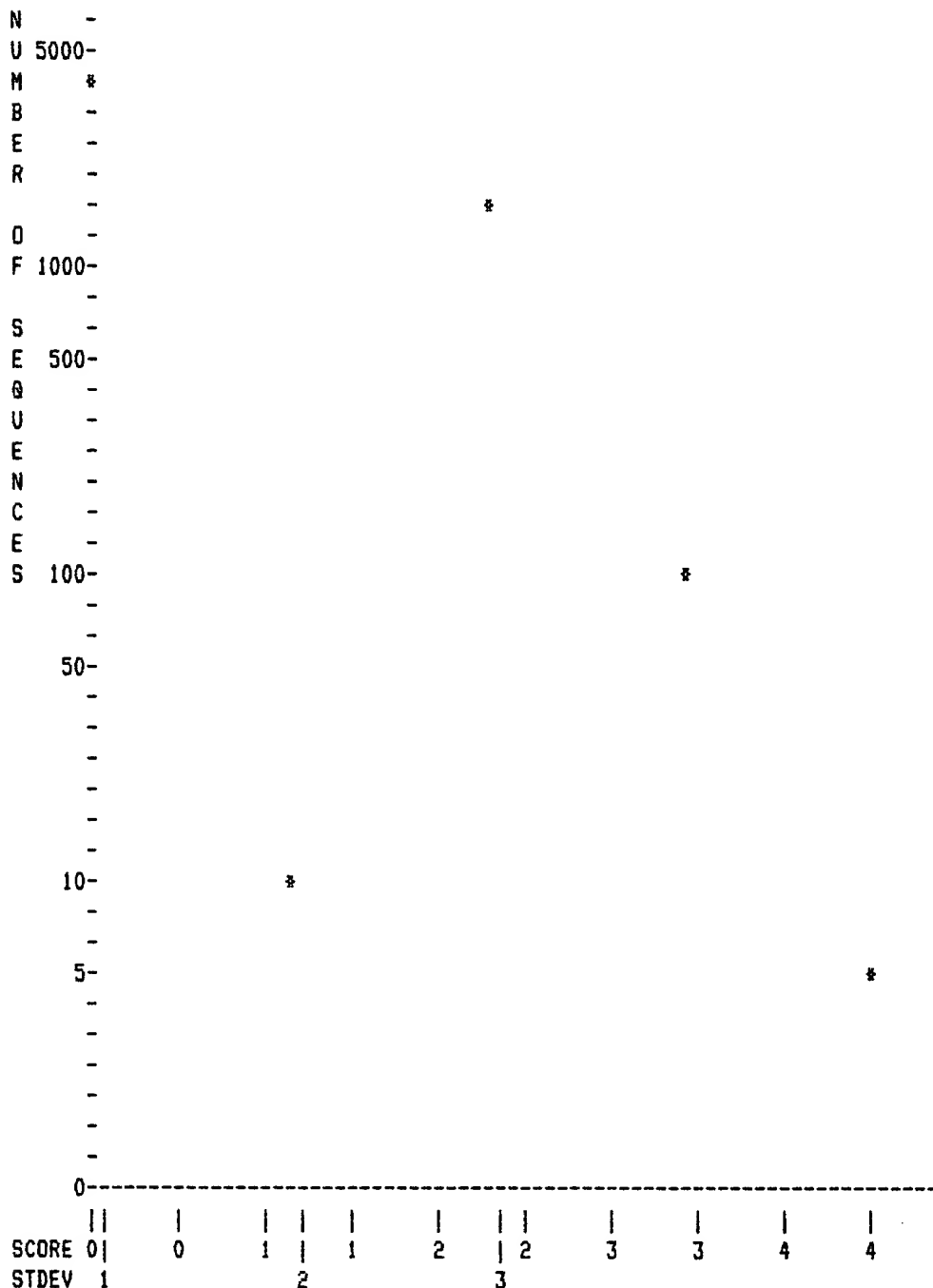
```
#####
#####
#####
#####
#####
##### User: stic7!shears
#####
##### Title: us-08-249-182-1.res
#####
##### stic7
#####
##### Printed: Wed 16:26 Sep 21, 1994
#####
##### Job number: MT661-7-87
#####
#####
#####
#####
#####
> 0 <
0| 0 IntelliGenetics
> 0 <
```

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-08-249-182-1.res made by on Wed 21 Sep 94 12:06:27-PDT.

Query sequence being compared: US-08-249-182-1 (1-5)
Number of sequences searched: 5543
Number of scores above cutoff: 1605

Results of the initial comparison of US-08-249-182-1 (1-5) with:
File : /home/shears/loring/loring*.pep



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	0.95
Times:	CPU	Total Elapsed	
	00:00:10.85	00:00:19.00	

Number of residues: 482836

Number of sequences searched: 3343
Number of scores above cutoff: 1605

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
1. R37443	Autotaxin peptide ATX 18.	6	4	4	4.23	0
2. R42697	p16 of nef peptide of HIV-1.	35	4	4	4.23	0
3. R42698	p17 of nef peptide of HIV-1.	65	4	4	4.23	0
4. R38893	Nef protein of HIV-1.	206	4	4	4.23	0
**** 3 standard deviations above mean ****						
5. R34888	Human TSH residues 61-75.	15	3	3	3.17	0
6. R30061	Human PMP sequence used to ra	16	3	3	3.17	0
7. P30110	Sequence of VP1 capsid prote	18	3	3	3.17	0
8. P90011	Synthetic peptide pol-1 corre	18	3	3	3.17	0
9. R36792	Prion protein region A subfra	19	3	3	3.17	0
10. R37300	E.coli shiga-like toxin segme	20	3	3	3.17	0
11. P30108	Sequence of VP1 capsid prote	20	3	3	3.17	0
12. P98462	Sequence of C. trachomatis se	22	3	3	3.17	0
13. P98458	Sequence of C. trachomatis se	22	3	3	3.17	0
14. P98454	Sequence of C. trachomatis se	22	3	3	3.17	0
15. P98450	Sequence of C. trachomatis se	22	3	3	3.17	0
16. P98446	Sequence of C. trachomatis se	22	3	3	3.17	0
17. P98442	Sequence of C. trachomatis se	22	3	3	3.17	0
18. P98466	Sequence of C. trachomatis se	22	3	3	3.17	0
19. R41085	HTLV-I and HTLV-II peptide I-	23	3	3	3.17	0
20. R11555	Native HIV gp160 peptide (2)	24	3	3	3.17	0
21. R30757	HIV discriminatory peptide, H	27	3	3	3.17	0
22. R38030	Ovine prion protein region E	30	3	3	3.17	0
23. R38027	Bovine prion protein region E	30	3	3	3.17	0
24. R31565	HTLV-II epitope, HTLV-II env-	30	3	3	3.17	0
25. R31547	HTLV-I epitope, HTLV-I env-4.	31	3	3	3.17	0
26. R31546	HTLV-I epitope, HTLV-I env-3.	31	3	3	3.17	0
27. P30122	Sequence of VP1 capsid protei	31	3	3	3.17	0
28. P30107	Sequence of VP1 capsid prote	31	3	3	3.17	0
29. P30106	Sequence of VP1 capsid prote	31	3	3	3.17	0
30. P90947	Peptide 240.	31	3	3	3.17	0
31. R39873	C peptide RV-C6, residues 119	34	3	3	3.17	0
32. R44497	Sequence of the HIV-1 epitope	35	3	3	3.17	0
33. R39893	Lipopeptide TPRV-C6.	37	3	3	3.17	0
34. R41092	HTLV-I and HTLV-II peptide II	40	3	3	3.17	0
35. P90948	Peptide 242.	41	3	3	3.17	0
36. R38503	P. aeruginosa pilin protein s	53	3	3	3.17	0
37. R33877	Polypeptide p380.LG comprisin	57	3	3	3.17	0
38. B42575	Abl heavy chain variable reg	101	3	3	3.17	0
39. R30763	Heavy chain variable domain o	120	3	3	3.17	0
40. R47104	Human NT-4 encode by genonic	132	3	3	3.17	0

1. US-08-249-182-1 (1-5)

R37443 Autotaxin peptide ATX 18.

ID R37443 standard; peptide; 6 AA.
AC R37443;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 18.

KW cell motility stimulating; cancer metastasis; antibody; detection;
 KW immunostains; disease outcome prediction; therapy choice;
 KW cancer therapy; crosslinked toxins.
 OS Synthetic.
 PN US7822043-A.
 PD 01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 18. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 6 AA;
 SQ 2 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 1 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 0 S; 0 T; 1 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 4.23
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 WHVAAN
 X X

2. US-08-249-182-1 (1-5)

R42697 p16 of nef peptide of HIV-1.

ID R42697 standard; protein; 35 AA.
 AC R42697;
 DT 10-NOV-1993 (first entry)
 DE p16 of nef peptide of HIV-1.
 KW AIDS; antibody; p25; gp110; gp41; assay; detection;
 KW immunity; vaccine.
 OS Human immunodeficiency virus-1.
 FH Key Location/Qualifiers
 FT Modified_site 35
 FT /note= "Cys(acetanidomethyl)"
 PN US5221610-A.
 PD 22-JUN-1993.
 PF 26-MAY-1988; 199143.
 PR 26-MAY-1988; US-199143.
 PR 04-SEP-1991; US-754300.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 PI Bahraoui EM, Chanaret S, Ferris S, Granier C, Montagnier L;
 PI Rietschoten JV, Rochat H, Sabatier JM;
 DR WPI; 93-213434/26.
 PT Diagnosis of HIV infection - by detecting HIV antibodies using
 PT antigenic polypeptide derived from nef protein of HIV-1
 PS Disclosure; Page 3; 15pp; English.
 CC The peptide is expressed in vivo in HIV infected patients before
 CC detectable ants. of p25, gp110 and gp41 are expressed. Thus, it

CC Can be used in assays for early detection of HIV.
 CC It can also be used to raise antibodies for use in detection,
 CC to induce cellular immunity or to raise neutralising antibodies
 CC that either inactivate the AIDS virus or reduce the viability of
 CC the virus in vivo or destroy infected cells.
 CC The peptide may be used in viral vaccines.
 SQ Sequence 35 AA;
 SQ 2 A; 4 R; 1 N; 3 D; 0 B; 1 C; 0 Q; 5 E; 0 Z; 1 G; 3 H;
 SQ 0 I; 3 L; 1 K; 1 M; 3 F; 2 P; 1 S; 0 T; 1 W; 1 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 4.23
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      WHVAR
      ||||
GMDDPEREVLEWRFD SRLAFHHVARELHPEYFKNC
      10      20      X      30

```

3. US-08-249-182-1 (1-5)

R42698 p17 of nef peptide of HIV-1.

ID R42698 standard; protein; 65 AA.
 AC R42698;
 DT 10-NOV-1993 (first entry)
 DE p17 of nef peptide of HIV-1.
 KW AIDS; antibody; p25; gp110; gp41; assay; detection;
 KW immunity; vaccine.
 OS Human immunodeficiency virus-1.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "Cys(acetamidomethyl)"
 FT Modified_site 65
 FT /note= "Cys(acetamidomethyl)"
 PN US5221610-A.
 PD 22-JUN-1993.
 PF 26-MAY-1988; 199143.
 PR 26-MAY-1988; US-199143.
 PR 04-SEP-1991; US-754300.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;
 PI Rietschoten JV, Rochat H, Sabatier JM;
 DR WPI; 93-213434/26.
 PT Diagnosis of HIV infection - by detecting HIV antibodies using
 PT antigenic polypeptide derived from nef protein of HIV-1
 PS Disclosure; Page 3; 15pp; English.
 CC The peptide is expressed in vivo in HIV infected patients before
 CC detectable ants. of p25, gp110 and gp41 are expressed. Thus, it
 CC can be used in assays for early detection of HIV.
 CC It can also be used to raise antibodies for use in detection,
 CC to induce cellular immunity or to raise neutralising antibodies
 CC that either inactivate the AIDS virus or reduce the viability of
 CC the virus in vivo or destroy infected cells.
 CC The peptide may be used in viral vaccines.
 SQ Sequence 65 AA;
 SQ 3 A; 4 R; 3 N; 4 D; 0 B; 2 C; 0 Q; 9 E; 0 Z; 2 G; 5 H;
 SQ 0 I; 7 L; 4 K; 1 M; 3 F; 5 P; 3 S; 1 T; 1 W; 2 Y; 6 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 4.23
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||

CYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREVLEWRFD SRLAFHHVARELHPEYFKNC

10 20 30 40 50 X 60

4. US-08-249-182-1 (1-5)

R38893 Nef protein of HIV-1.

ID R38893 standard; Protein; 206 AA.
AC R38893;
DT 10-NOV-1993 (first entry)
DE Nef protein of HIV-1.
KW AIDS; antibody; p25; gp110; gp41; assay; detection;
KW immunity; vaccine.
OS Human immunodeficiency virus-1.
PN US5221610-A.
PD 22-JUN-1993.
PF 26-MAY-1988; 199143.
PR 26-MAY-1988; US-199143.
PR 04-SEP-1991; US-754300.
PA (INRM) INST NAT SANTE & RECH MEDICALE.
PA (INSP) INST PASTEUR.
PI Bahraoui EM, Chanaret S, Ferris S, Granier C, Montagnier L;
PI Rietschoten JV, Rochat H, Sabatier JM;
DR WPI; 93-213434/26.
PT Diagnosis of HIV infection - by detecting HIV antibodies using
PT antigenic polypeptide derived from nef protein of HIV-1
PS Disclosure; Fig 2; 15pp; English.
CC The nef protein comprises peptides which are expressed in vivo in HIV
CC infected patients before detectable ants. of p25, gp110 and gp41 are
CC expressed. Thus, they can be used in assays for early detection of HIV.
CC They can also be used to raise antibodies for use in detection,
CC to induce cellular immunity or to raise neutralising antibodies
CC that either inactivate the AIDS virus or reduce the viability of
CC the virus in vivo or destroy infected cells.
CC The peptides may be used in viral vaccines.
SQ Sequence 206 AA;
SQ 17 A; 13 R; 6 N; 10 D; 0 B; 3 C; 6 Q; 19 E; 0 Z; 16 G; 9 H;
SQ 4 I; 17 L; 10 K; 4 M; 7 F; 15 P; 11 S; 10 T; 7 W; 7 Y; 15 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 4.23
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||

CYKLVPEPDKVEEANKGENTSLHPVSLHGMDPEREVLEWRFD SRLAFHHVARELHPEYFKNC

150 160 170 180 190 X X 200

5. US-08-249-182-1 (1-5)

R34888 Human TSH residues 61-75.

ID R34888 standard; peptide; 15 AA.
AC R34888;
DT 23-JUL-1993 (first entry)
DE Human TSH residues 61-75.
KW Thyroid stimulating hormone; beta subunit; inhibitor; cAMP;
KW autoantibodies; Graves' disease.
OS Synthetic.
PN US5196513-A.

PF 05-SEP-1989; 403564.
 PR 05-SEP-1989; US-403564.
 PR 25-JUL-1991; US-736030.
 PA (MAYO-) MAYO FOUNDATION.
 PI McCormick DJ, Morris JC, Ryan RJ;
 DR WPI; 93-116855/14.
 PT Synthetic peptide derived from beta-sub-unit residues 101-112 of
 PT TSH - inhibits binding of human thyroid stimulating hormone to
 PT thyroid membrane, useful in immuno-therapy and immuno-diagnosis
 PT e.g. of Graves disease
 PS Disclosure; Page 12; 16pp; English.
 CC The sequence represents amino acid residues 61-75 of the beta subunit
 CC of human thyroid stimulating hormone (hTSH). This synthetic peptide
 CC is capable of inhibiting binding of TSH to human thyroid membranes,
 CC inhibiting TSH-mediated cAMP generation and inhibiting the
 CC stimulatory effect of thyroid stimulating autoantibodies. The
 CC peptide may be used as a diagnostic and therapeutic reagent for
 CC thyroid diseases such as Graves' disease. The peptide may also be
 CC used to produce antibodies which can be used to measure thyroid
 CC stimulating autoantibody levels or to therapeutically neutralise
 CC these autoantibodies.
 CC See also R34882-92.
 SQ Sequence 15 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 1 C; 0 Q; 1 E; 0 Z; 1 G; 1 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 3 P; 0 S; 1 T; 0 W; 1 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:46-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 TVEIPGCPLHVAPYF
 10 X

6. US-08-249-182-1 (1-5)

R30061 Human PMP sequence used to raise antibodies.

ID R30061 standard; peptide; 16 AA.
 AC R30061;
 DT 29-APR-1993 (first entry)
 DE Human PMP sequence used to raise antibodies.
 KW Charcot-Marie-Tooth disease; hypertrophic neuropathy; autosomal;
 KW CTM1A; hybridisation; PMP; peripheral myelin protein; Schwann.
 OS Homo sapiens.
 PN W09221694-A.
 PD 10-DEC-1992.
 PF 05-JUN-1992; U04833.
 PR 06-JUN-1991; US-711615.
 PR 06-MAY-1992; US-879623.
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Cardenas OS, De Oca-luna RM, Deleon M, Lupski JR, Patel PI;
 PI Shooter EM, Snipes GJ, Suter U, Welcher A;
 DR WPI; 92-433607/52.
 PT Human peripheral myelin protein and nucleic acid sequence - used
 PT for diagnosis and treatment of hypertrophic neuropathies e.g.
 PT autosomal dominant Charcot-Marie-Tooth disease
 PS Disclosure; Fig 20; 93pp; English.
 CC The sequence is that of a fragment of human PMP, a peripheral myelin
 CC protein characterised in that it is expressed predominantly in
 CC peripheral Schwann cells, has a mol. wt. of 20 kD and has

CC substantial sequence homology to a PNP 0000. From human peripheral
 CC nervous tissue. The peptide was used to raise antibodies which may
 CC be used in diagnosis of Charcot-Marie-Tooth type 1 disease
 CC (CMT1A), which is a common peripheral neuropathy in humans. The
 CC DNA based test can be performed on DNA isolated from peripheral blood
 CC at less than one tenth of the cost of present diagnostic methods, and
 CC the disease may be detected early at the genetic level without the need
 CC for symptoms.
 CC See also R30056-60.
 SQ Sequence 16 AA;
 SQ 0 A; 2 R; 2 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 2 S; 0 T; 1 W; 3 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:32-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 YRVRHSEWHVNNDSYS
 10 X

7. US-08-249-182-1 (1-5)
 P30110 Sequence of VP1 capsid protein residues 141-160 f

ID P30110 standard; Protein; 18 AA.
 AC P30110;
 DT 03-APR-1992 (first entry)
 DE Sequence of VP1 capsid protein residues 141-160 from the amino-
 DE terminus, FMDV, Tubingen type 0, subtype 1, strain Kaufbeuren.
 KW Antigen; Picornavirus; capsid protein; antibody; detection;
 KW vaccine; diagnosis.
 OS Foot and mouth disease virus.
 PN W08303547-A.
 PD 27-OCT-1983.
 PF 14-APR-1983; 002644.
 PR 14-APR-1982; US-368308.
 PR 25-MAR-1983; US-478847.
 PR 20-SEP-1984; US-653475.
 PR 18-DEC-1984; US-682819.
 PA (BITT/) BITTLE J L.
 PA (SCRI-) SCRIPPS CLINIC & RE.
 PI Bittle JL, Lerner RA.
 DR WPI; 83-807942/44.
 PT Antigenic peptide(s) corresp. to picornavirus capsid protein -
 PT useful in prodn. of vaccines and in diagnostic tests
 PS Disclosure; Page 14; 90pp; English.
 CC The peptides of the invention corresp. to a region on the antigenic
 CC Picornavirus capsid protein. The capsid protein FMDV VP1 or polio
 CC virus VP1. When linked to carriers the peptides are immunogenic.
 CC Dose is 20 ug-2mg peptide for inoculations.
 SQ Sequence 18 AA;
 SQ 2 A; 3 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 3 G; 0 H;
 SQ 0 I; 2 L; 0 K; 0 M; 1 F; 2 P; 2 S; 0 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:59-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||

8. US-08-249-182-1 (1-5)

P90011 Synthetic peptide pol-1 corresp. to residues 822-8

ID P90011 standard; protein; 18 AA.
 AC P90011;
 DT 1-NOV-1989 (first entry)
 DE Synthetic peptide pol-1 corresp. to residues 822-838 of
 DE polymerase gene from hepatitis B ayw genome
 KW Hepatitis B virus peptide; determines antibody; synthetic;
 KW polymerase.
 OS Hepatitis virus
 PN W08904964-A.
 PD 01-JUN-1989.
 PF 15-NOV-1988; U04076.
 PR 16-NOV-1987; US-120979.
 PA (FOXC) Fox Chase Cancer CE.
 PI Feitelson M, Blumberg BS, Millman I;
 DR WPI; 89-178459/24.
 PT Immunoassay for antibody to hepatitis virus DNA polymerase
 PT - representing early marker of hepatitis infections
 PS Claim 3; page 30; 34pp; English.
 CC Synthetic peptide, pol-1, corresp. to amino acid residues
 CC 822-838 of the polymerase gene from hepatitis B ayw genome.
 CC Used for raising antibodies to determine hepatitis infection
 CC at an early stage, and to reduce the incidence of post-transfusion
 CC hepatitis B.
 SQ Sequence 18 AA;
 SQ 2 A; 2 R; 0 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 1 L; 0 K; 0 M; 1 F; 4 P; 1 S; 0 T; 1 W; 0 Y; 3 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:56:48-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 PVRVHFASPLHVAWRPPC
 10 X

9. US-08-249-182-1 (1-5)

R36792 Prion protein region A subfragment #1.

ID R36792 standard; protein; 19 AA.
 AC R36792;
 DT 14-OCT-1993 (first entry)
 DE Prion protein region A subfragment #1.
 KW Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
 KW FSa; FSb; subfragment; antibody; treatment; spongiform encephalopathy;
 KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
 KW immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;
 KW resistance.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1
 FT /note= "One or more residues or may be absent"
 FT Misc_difference 2
 FT /note= "May be absent"
 FT Misc_difference 3
 FT /note= "May be absent"

FT /note= "May be absent"
 FT Misc_difference 5
 FT /note= "May be absent"
 FT Misc_difference 15
 FT /note= "May be absent"
 FT Misc_difference 16
 FT /note= "May be absent"
 FT Misc_difference 17
 FT /note= "May be absent"
 FT Misc_difference 18
 FT /note= "May be absent"
 FT Misc_difference 19
 FT /note= "One or more residue or may be absent"
 PN W09311155-A.
 PD 10-JUN-1993.
 PF 03-DEC-1992; G02246.
 PR 03-DEC-1991; GB-025747.
 PR 10-JUL-1992; GB-014663.
 PA (PROT-) PROTEUS MOLECULAR DESIGN LTD.
 PI Fishleigh RV, Mee RP, Robson B;
 DR WPI; 93-196994/24.
 PT New polypeptide(s) contg. antigenic site of prion protein -
 PT useful for treatment and diagnosis of mammalian encephalopathies
 PT e.g. Creutzfeld-Jacob disease and kuru
 PS Claim 7; Page 63; 82pp; English.
 CC The sequences given in R36792-95 represent polypeptide subfragments
 CC derived from an antigenic site, region A, of a prion protein. Prion
 CC proteins comprise six regions of interest (A-F), and two related frame
 CC shift peptides sequences caused by a repeating section in region E
 CC having a nucleic acid coding sequence frame shift mutation of +1 (FSa)
 CC or -1 (FSb). The full length peptide region (see also R36786-89 and
 CC these subfragments, and antibodies raised against these, may be used
 CC to treat or prevent spongiform encephalopathy in humans, sheep or
 CC cattle. They can be used to block cellular binding and aggregation
 CC of prion proteins and to stimulate the mammalian immune system. These
 CC peptides may be used to distinguish between the normal form of prion
 CC protein (PrPc) and the scrapie-associated form (PrPsc). These
 CC peptides may include rare or synthetic amino acids or a ratio-inverso
 CC peptide modification to improve resistance to enzymatic degradation.
 SQ Sequence 19 AA;
 SQ 6 A; 0 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 6 G; 1 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 0 Y; 3 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:15-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 XHVAGAAAAGAVVGGGCGX
 X X 10

10. US-08-249-182-1 (1-5)
 R37300 E.coli shiga-like toxin segment.

ID R37300 standard; Protein; 20 AA.
 AC R37300;
 DT 13-SEP-1993 (first entry)
 DE E.coli shiga-like toxin segment.
 KW Type I ribosome-inactivating protein; ricin; monordin;
 KW immunoconjugate; autoimmune disease; cell killing; toxin;

KW human engineered antibody; variable region; light chain;
 KW cell targetting; chimeric antibody; SLT.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT Disulfide_bond 1..20
 FT /note= "intervening loop includes protease
 FT sensitive amino acid sequence"
 PN W09309130-A.
 PD 13-MAY-1993.
 PF 04-NOV-1992; U09487.
 PR 04-NOV-1991; US-787567.
 PR 19-JUN-1992; US-901707.
 PA (XDMA) XDMA CORP.
 PI Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP.
 DR WPI; 93-167617/20.
 PT Analogues of type I ribosome inactivating protein - useful as
 PT cytotoxic agents, immuno toxins for treating auto immune diseases,
 PT cancer, graft versus host disease and selective cell killing in-vivo
 PS Example 10; Page 114; 163pp; English.
 CC The invention covers analogues of the plant type I RIP gelonin
 CC which have a non-naturally occurring Cys residue in a position
 CC which enables the analogue to be conjugated via a disulphide
 CC linkage to a molecule which specifically binds to a target cell.
 CC Pref. target-cell binding molecules are antibodies or their
 CC fragments, esp. human engineered M65 antibody fragments. Fusion
 CC constructs were assembled that included a natural sequence gelonin
 CC gene fused to an M65 truncated heavy chain gene or an M65 light
 CC chain (kappa) gene. A DNA linker encoding a peptide segment of the
 CC E.coli shiga-like toxin was inserted between the gelonin gene and
 CC the Ab gene. The resulting immunoconjugates can be used as cytotoxic
 CC therapeutic agents.
 SQ Sequence 20 AA;
 SQ 3 A; 2 R; 0 N; 1 D; 0 B; 2 C; 0 Q; 1 E; 0 Z; 0 G; 3 H;
 SQ 0 I; 0 L; 0 K; 2 M; 1 F; 1 P; 3 S; 0 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:09-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 CHHHASRVARMASDEFPSMC
 X 10 20

11. US-08-249-182-1 (1-5).

P30108 Sequence of VP1 capsid protein residues 141-160 f

ID P30108 standard; Peptide; 20 AA.
 AC P30108;
 DT 03-APR-1992 (first entry)
 DE Sequence of VP1 capsid protein residues 141-160 from the amino-
 DE terminus, FMDV, Tubingen type O, subtype 1, strain Kaufbeuren.
 KW Antigen; Picornavirus; capsid protein; antibody; detection;
 KW vaccine; diagnosis.
 OS Foot and mouth disease virus.
 PN W08303547-A.
 PD 27-OCT-1983.
 PF 14-APR-1983; 002644.
 PR 14-APR-1982; US-368308.
 PR 25-MAR-1983; US-478847.
 PR 20-SEP-1984; US-653475.
 PR 18-DEC-1984; US-682819.
 PA (BITT/) BITTLE J L.
 PA (BITT/) BITTLE J L.

PI Bittle JL, Lerner RA.
 DR WPI: 83-807942/44.
 PT Antigenic peptide(s) corresp. to picornavirus capsid protein -
 PT useful in prodn. of vaccines and in diagnostic tests
 PS Disclosure: Page 14; 90pp; English.
 CC The peptides of the invention corresp. to a region on the antigenic
 CC Picornavirus capsid protein. The capsid protein FMDV VP1 or polio
 CC virus VP1. When linked to carriers the peptides are immunogenic.
 CC Dose is 20 ug-2mg peptide for inoculations.
 SQ Sequence 20 AA;
 SQ 2 A; 2 R; 1 N; 1 D; 0 B; 0 C; 2 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 4 L; 1 K; 0 M; 0 F; 2 P; 0 S; 1 T; 0 W; 0 Y; 3 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:59-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 VPNLRGDLQVLAKKVARTLP
 10 X X 20

12. US-08-249-182-1 (1-5)

P98462 Sequence of *C. trachomatis* serovar L3 major outer

ID P98462 standard; Protein; 22 AA.
 AC P98462;
 DT 06-MAR-1992 (first entry)
 DE Sequence of *C. trachomatis* serovar L3 major outer membrane protein (MOMP)
 DE variable domain (VD) gene L3-VDIV base pairs 487-552
 KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay; ss.
 OS *Chlamydia trachomatis*.
 FH Key Location/Qualifiers
 FT CDS 1..66
 FT /*tag= a
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI: 89-339697/46.
 DR P-PSDB; P98467.
 PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure: Fig 19; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 22 AA;
 SQ 3 A; 1 R; 2 N; 2 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 3 P; 1 S; 4 T; 0 W; 0 Y; 2 V;

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      WHVAR
      |||
AAPTTS DVEGLQNDPTTNVARP
      10      X 20
  
```

13. US-08-249-182-1 (1-5)

P98458 Sequence of *C. trachomatis* serovar J major outer m

ID P98458 standard; Protein; 22 AA.
 AC P98458;
 DT 06-MAR-1992 (first entry)
 DE Sequence of *C. trachomatis* serovar J major outer membrane protein (MOMP)
 DE variable domain (VD) J-VDI encoded by base pairs 256-321
 KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay.
 OS *Chlamydia trachomatis*.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97093.
 PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 17; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 22 AA;
 SQ 4 A; 1 R; 2 N; 2 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 3 P; 1 S; 4 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:44-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      WHVAR
      |||
AAPTTS DVAGLQNDPTTNVARP
      10      X 20
  
```

14. US-08-249-182-1 (1-5)

P98454 Sequence of *C. trachomatis* serovar I major outer m

ID P98454 standard; Protein; 22 AA.
 AC P98454;
 DT 06-MAR-1992 (first entry)
 DE Sequence of *C. trachomatis* serovar I major outer membrane protein (MOMP)
 DE variable domain (VD) I-VDI encoded by base pairs 256-321
 KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay.
 OS *Chlamydia trachomatis*.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97089.
 PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 16; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 22 AA;
 SQ 4 A; 1 R; 2 N; 2 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 1 L; 1 K; 0 M; 0 F; 3 P; 0 S; 4 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:44-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

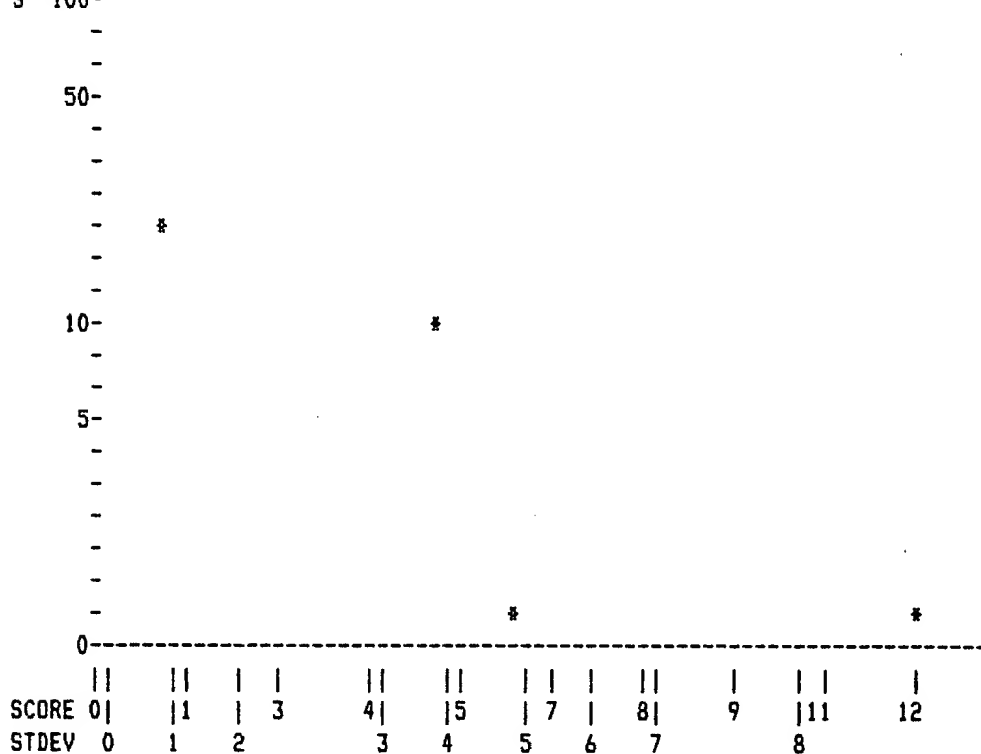
```

      X  X
      WHVAR
      |||
AAPTTKDVAGLENDPTTNVARP
    10      X 20
  
```

15. US-08-249-182-1 (1-5)

P98450 Sequence of *C. trachomatis* serovar H major outer m

ID P98450 standard; Protein; 22 AA.
 AC P98450;
 DT 06-MAR-1992 (first entry)
 DE Sequence of *C. trachomatis* serovar H major outer membrane protein (MOMP)
 DE variable domain (VD) H-VDI encoded by base pairs 256-321
 KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay.
 OS *Chlamydia trachomatis*.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97085.



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.21

Times:	CPU	Total Elapsed
	00:00:08.91	00:00:08.00

Number of residues:	482836
Number of sequences searched:	5543
Number of scores above cutoff:	2815

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
1. R37452	Autotaxin peptide ATX 102.	12	12	12	9.06 0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure: Fig 15; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 22 AA:
 SQ 5 A; 1 R; 3 N; 3 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 1 K; 0 M; 0 F; 3 P; 0 S; 3 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:44-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.17
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||
 AAPTTNDAADLQNDPKTNVARP
 10 X 20

> 0 <
 0| 10 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-08-249-182-10.res made by on Wed 21 Sep 94 12:05:15-PDT.

Query sequence being compared:US-08-249-182-10 (1-12)
 Number of sequences searched: 5543
 Number of scores above cutoff: 2815

Results of the initial comparison of US-08-249-182-10 (1-12) with:
 File : /hone/shears/loring/lorin*.pep

10000-
 -
 N -
 U 5000-
 M -
 B *
 E - *
 R -
 -
 O -
 F 1000-
 -
 S - *
 E 500-
 Q -
 U -
 E -
 N -
 C -
 E - *

2. R08086	Feline T-cell lymphotropic l	57	6	6	4.12	0
**** 3 standard deviations above mean ****						
3. P82053	Outer membrane protein F of P	350	5	5	3.30	0
4. R41746	MN protein.	429	5	6	3.30	0
5. R40856	43kd regression associated an	453	5	5	3.30	0
6. R40916	Sequence of a CD26 fragment l	593	5	5	3.30	0
7. R40909	Sequence encoded by human CD2	766	5	5	3.30	0
8. P50287	Sequence encoded by hepatitis	854	5	5	3.30	0
9. P50231	Sequence encoded by partial s	993	5	5	3.30	0
10. P98500	Partial sequence encoded by n	1091	5	5	3.30	0
11. R12141	Enteroviral polypeptide.	2185	5	5	3.30	0
**** 2 standard deviations above mean ****						
12. P30304	Sequence corresp. to a protei	12	4	4	2.47	0
13. R41129	HCV peptide XXc-1 (aa 383-404	24	4	5	2.47	0
14. R39589	Ribonucleoprotein snRNP-U1 pr	24	4	4	2.47	0
15. R25061	Thyrotrope reacting peptide 1	24	4	4	2.47	0
16. R10304	HTLV-1 antibody epitope (v).	24	4	4	2.47	0
17. R10305	HTLV-1 antibody epitope (vi).	24	4	4	2.47	0
18. R41130	HCV peptide XXc-2 (aa 393-416	26	4	5	2.47	0
19. R41128	HCV peptide XXc (aa 383-416;	36	4	5	2.47	0
20. R40189	Sequence of peptide construct	37	4	4	2.47	0
21. R30905	Amino acids 239-278 of human	39	4	4	2.47	0
22. R10317	Bovine BMP - exon 3.	41	4	5	2.47	0
23. P80685	Peptide 141 from the HIV p25	45	4	4	2.47	0
24. R33885	Polypeptide p380-JH1 comprisi	57	4	5	2.47	0
25. R29584	Rana NT-4 protein.	57	4	4	2.47	0
26. R14456	HIV-1 hxb2 gag 213-273 (2).	69	4	4	2.47	0
27. R14006	HIV-1 hxb2 gag 213-273 (1).	69	4	4	2.47	0
28. R40847	Metalloproteinase-I.	84	4	4	2.47	0
29. R11715	HTLV-1 env. gp21 epitope enco	106	4	4	2.47	0
30. R42804	RSV19 heavy chain variable re	116	4	5	2.47	0
31. R42802	RSV19 heavy chain variable re	116	4	5	2.47	0
32. R24807	RSV19 VH.	116	4	5	2.47	0
33. R07322	VH domain of antibody D again	116	4	4	2.47	0
34. G45722	anti-glycoprotein H monoclon	117	4	4	2.47	0
35. R39464	MAB BW 2128 heavy chain V-reg	119	4	4	2.47	0
36. R29710	p16 gag protein from hTLR.	123	4	5	2.47	0
37. R29011	p146-h1 protein product.	134	4	4	2.47	0
38. R28669	p12-h2.	135	4	4	2.47	0
39. R12360	Heavy chain variable region o	137	4	4	2.47	0
40. R11384	Variable gamma heavy chain of	140	4	4	2.47	0

1. US-08-249-182-10 (1-12)

R37452 Autotaxin peptide ATX 102.

ID R37452 standard; peptide; 12 AA.
AC R37452;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 102.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 102. It may be used to

CC false anti autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 12 AA;
 SQ 0 A; 1 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 1 H;
 SQ 1 I; 2 L; 0 K; 0 M; 2 F; 0 P; 1 S; 1 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 12 Optimized Score = 12 Significance = 9.06
 Residue Identity = 100% Matches = 12 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 |||||
 DIEHLTSLDFFR
 X 10 X

2. US-08-249-182-10 (1-12)
 R08086 Feline T-cell lymphotropic lentivirus of clone R5

ID R08086 standard; protein; 57 AA.
 AC R08086;
 DT 26-FEB-1991 (first entry)
 DE Feline T-cell lymphotropic lentivirus of clone R5XCXL1.
 KW Feline T-cell lymphotropic lentivirus; FIV; R5XCXL1; antibodies;
 KW vaccines.
 OS Feline T-cell lymphotropic lentivirus 2428 (Pentaluma).
 PN W09013573-A.
 PD 15-NOV-1990.
 PF 30-APR-1990; U02338.
 PR 08-MAY-1989; US-348784.
 PR 08-DEC-1989; US-447810.
 PA (IDEX-) IDEXX CORP.
 PI Anderson PR, Oconnor TP, Tonelli QJ;
 DR WPI; 90-361429/48.
 DR N-PSDB; Q06655.
 PT Feline T-cell lympho-tropic lentivirus poly-peptide(s) - used for
 PT specific detection of FIV antibodies, prodn. of antibodies and in
 PT vaccines
 PS Disclosure; Fig 5(c); 37pp; English.
 CC The amino acid sequence shows homology with the envelope gene of
 CC equine infectious anemia virus, a lentivirus, immunologically
 CC closely related to FIV.
 CC Strain R5X has been deposited ATCC 67939.
 CC See also Q06653-55 and R08094-96.
 SQ Sequence 57 AA;
 SQ 1 A; 3 R; 5 N; 2 D; 0 B; 2 C; 2 Q; 2 E; 0 Z; 1 G; 1 H;
 SQ 2 I; 6 L; 0 K; 0 M; 7 F; 3 P; 6 S; 3 T; 2 W; 0 Y; 7 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:10-PDT using FindSeq

Initial Score = 6 Optimized Score = 6 Significance = 4.12
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 ||| ||
 ATNSQSCENVFSSXLPNGHNTVWEQVVFLSIFWDLFXVRNCVLTSLFPFRDLVIF
 10 20 30 40 50 X

3. US-08-249-182-10 (1-12)

P82053 Outer membrane protein F of *Pseudomonas aeruginosa*

ID P82053 standard; protein; 350 AA.
 AC P82053;
 DT 22-OCT-1990 (first entry)
 DE Outer membrane protein F of *Pseudomonas aeruginosa*.
 KW Outer membrane protein F; OMPF; vaccination; antibodies.
 KW
 OS *Pseudomonas aeruginosa*.
 PN DE3718591-A.
 PD 15-DEC-1987.
 PF 03-JUN-1987; 718591.
 PR 03-JUN-1987; DE-718591.
 PA (BEHW) Behringwerke AG.
 PI Domdey H, Lottspeich F, von Specht B-U, Duchene M;
 DR WPI; 88-361619/51.
 DR N-PSDB; N82023.
 PT New outer membrane protein F of *Pseudomonas aeruginosa* -
 PT DNA sequences encoding it and derived antibodies, useful for
 PT vaccination and diagnosis.
 PS Disclosure; p; German.
 CC The protein is isolated from the OMP of *P. aeruginosa* serotype 6
 CC ATCC 33354) and purified by HPLC. The amino-terminal and trypsin
 CC fragments are sequenced and a series of oligonucleotide probes
 CC constructed corresponding to the established sequences. These probes are
 CC used to screen a gene bank of 15-20 kb fragments of genomic DNA in
 CC lambda EMBL 3. One positive clone includes a 15 kb insert contg. the
 CC protein gene, which can be isolated as a 2.5 kb PstI fragment. This
 CC fragment cannot be cloned into a high copy no. vector because of the
 CC toxicity of the gene prod., so is subcloned as two fragments with an
 CC overlapping region of about 500bp. Ab's are raised by usual immunisation
 CC or cell-fusion procedures. The Ab's are useful in diagnosis.
 SQ Sequence 350 AA;
 SQ 37 A; 15 R; 26 N; 26 D; 0 B; 4 C; 11 Q; 21 E; 0 Z; 37 G; 6 H;
 SQ 9 I; 19 L; 19 K; 7 M; 13 F; 10 P; 21 S; 18 T; 1 W; 16 Y; 34 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:00-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR

||||

RYFTDSVRNMKNADLYGGSIGYFLTDDVELALSYGEYHDVRGTYETGNKKVHGNTSLDAIYHFGTPGVGLR
 40 50 60 70 80 90 100

PYVSAGLAHQNITNINSDSQGRQMTMANIGAGLKYFTE
 110 120 130 140

4. US-08-249-182-10 (1-12)

R41746 MN protein.

ID R41746 standard; Protein; 429 AA.
 AC R41746;
 DT 25-MAR-1994 (first entry)
 DE MN protein.
 KW MN; endogenous; MaTu; quasi-viral agent; human; mammary tumour; prion;
 KW classical virus; slow virus; exogenous MX; p58X; cytoplasmic antigen;
 KW conservative; HeLa cell; twin protein; p54/58N; cell surface; nucleus;
 KW monoclonal antibody; MAb M75; neoplasm; pre-neoplastic disease;
 KW vaccine.

US Homo sapiens.
 PN W09318152-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; U02024.
 PR 11-MAR-1992; CS-000709.
 PR 21-OCT-1992; US-964589.
 PA (CIBA) CIBA CORNING DIAGNOSTICS CORP.
 PA (VIRO-) INST VIROLOGY.
 PI Pastorek J, Pastorekova S, Zavada J;
 DR WPI; 93-303466/38.
 DR N-PSDB; Q48456.
 PT New MN gene and polypeptide(s) - used in diagnosis, prognosis and
 PT therapy of neoplastic and/or pre-neoplastic disease
 PS Claim 7; Fig 1; 72pp; English.
 CC This sequence is encoded by the intronless MN gene which is a cellular
 CC gene which is the endogenous component of the MaTu agent. MaTu is a
 CC novel quasi-viral agent with rather unusual properties. It is
 CC presumably derived from a human mammary tumour. In some aspects it
 CC resembles classical viruses, whereas in other respects it resembles
 CC "slow" viruses (prions), and in still other aspects it is different
 CC from both classes of viruses. MaTu is a two component sytem. One
 CC part of the complex, exogenous MX, is transmissible, and is manifest
 CC by a protein, p58X, which is a cytoplasmic antigen which reacts with
 CC some natural sera, of humans and of various animals. The other
 CC component, MN, is endogenous to human cells. MN is a cellular gene
 CC showing very little homology with known DNA sequences. It is rather
 CC conservative and present as a single copy in the chromosomal DNA of
 CC various vertebrates. MN is manifest in HeLa cells by a twin protein
 CC p54/58N, that is localised on the cell surface and in the nucleus.
 CC Immunoblots using a monoclonal antibody reactive with p54/58N (MAb M75)
 CC revels two bands at 54 kD and 58 kD. These two bands may correspond to
 CC one type of protein that differs by glycosylation pattern or by how it
 CC is processed. The expression of the MN gene is strongly correlated
 CC with tumourigenicity. MN products can be used in can be used in
 CC diagnostic and/or prognostic assays for neoplastic and/or pre-
 CC neoplastic disease. MN polypeptides, produced recombinantly by
 CC unicellular hosts, can also be used for antibody production and in
 CC vaccines for inducing protective immunity against neoplastic disease
 CC and a dampening effect upon tumourigenic activity.
 SQ Sequence 429 AA;
 SQ 41 A; 43 R; 8 N; 17 D; 0 B; 8 C; 21 Q; 28 E; 0 Z; 33 G; 13 H;
 SQ 13 I; 40 L; 10 K; 6 M; 11 F; 39 P; 40 S; 16 T; 13 W; 7 Y; 22 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:31-PDT using FindSeq

Initial Score = 5 Optimized Score = 6 Significance = 3.30
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || | || |
VDEALGRPGGLAVLAPFWRARRKKTVSYEQLLSRLEEIAEEGSETQVPLDISALLPSDFSRYFQYEGSLTT
    230      240      250      260      270 X 280 X 290

PPCAQGVIVTVFNQTVMLSAKQLHTLSDTLWGPGDSRLQL
    300      310      320      330
  
```

5. US-08-249-182-10 (1-12)

R40856 43kd regression associated antigen.

ID R40856 standard; Protein; 453 AA.
 AC R40856;
 DT 07-MAR-1994 (first entry)
 DE 43kd regression associated antigen.
 KW Regression associated antigen; tumour; immunotherapy;

KW anti-idiotypic antibodies; antibodies; tumour regression.
 OS Mycoplasma hyorhinitis.
 FH Key Location/Qualifiers
 FT Misc_difference 80
 FT /note= "Tryptophan encoded by TGA, normal in
 FT Mycoplasma hyorhinitis."
 FT Misc_difference 124
 FT /note= "Tryptophan encoded by TGA, normal in
 FT Mycoplasma hyorhinitis."
 FT Misc_difference 165
 FT /note= "Tryptophan encoded by TGA, normal in
 FT Mycoplasma hyorhinitis."
 FT Misc_difference 344
 FT /note= "Tryptophan encoded by TGA, normal in
 FT Mycoplasma hyorhinitis."
 PN US5242823-A.
 PD 07-SEP-1993.
 PF 07-MAR-1986; 837494.
 PR 07-MAR-1986; US-837494.
 PR 16-SEP-1987; US-097910.
 PR 11-DEC-1987; US-131815.
 PR 04-JAN-1988; US-138923.
 PR 16-MAR-1990; US-474730.
 PR 02-OCT-1992; US-956546.
 PA (ITGE-) INT GENETIC ENG INC.
 PI Fareed GC, Ghosh-dastidar P, Jar-how L, Sen A;
 DR WPI; 93-295229/37.
 DR N-PSDB; 047816.
 PT DNA encoding a regression-associated antigen from M. hyorhinitis -
 PT is used to obtain prods. for diagnosis, localisation and therapy
 PT of tumours
 PS Disclosure; Figure 3; 40pp; English.
 CC Regression associated antigens (RAA's) are identified in material
 CC from neoplastic cells by their immunological reactivity with
 CC regression associated antibodies from the serum of patients
 CC diagnosed as undergoing regression of a tumour. RAA's can be used
 CC for tumour immunotherapy and for producing and purifying antibodies
 CC which can be used for tumour diagnosis, localisation and therapy.
 CC The antibodies can also be used for the production of
 CC anti-idiotypic antibodies which can also be used in immunotherapy.
 SQ Sequence 453 AA;
 SQ 34 A; 6 R; 35 N; 33 D; 0 B; 1 C; 14 Q; 22 E; 0 Z; 30 G; 4 H;
 SQ 35 I; 33 L; 49 K; 2 M; 30 F; 10 P; 32 S; 39 T; 4 W; 11 Y; 29 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:26-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || ||
 FDNSEFVKDRQAEIEKAKNFDFNTVLLTAGCTVQDKSFNQSIWEAVLEHYDQIEKTTNLDKRVSGETNNQSELI
 40 50 60 70 80 90 100 110
 GKYKNFLNGKNVWILTGFQGGQEFKFLKQDSNGKKYS
 120 130 140 150

6. US-08-249-182-10 (1-12)

R40916 Sequence of a CD26 fragment lacking a portion of t

ID R40916 standard; Protein; 593 AA.
 AC R40916;
 DT 05-FEB-1994 (first entry)
 DE Sequence of a CD26 fragment lacking a portion of the carboxy

KW Human T cell activation antigen; monoclonal antibody Tai; CD26.
 OS Synthetic.
 PN W09316102-A.
 PD 19-AUG-1993.
 PF 09-APR-1992; U02892.
 PR 06-FEB-1992; US-832211.
 PA (DAND) DANA FARBER CANCER INST INC.
 PI Morimoto C, Schlossman SF, Tanaka T;
 DR WPI; 93-272827/34.
 PT Polypeptide fragments of CD26 - are capable of disrupting binding
 PT of CD45 and CD26 and thus interfering with T-cell activation
 PS Example; Pages 46-48; 73pp; English.
 CC C26 is a human T cell activation antigen originally identified by
 CC its reactivity with the MAb Tai. C26 cDNA library was constructed
 CC from human PHA-activated T cells using the CDM7 vector.
 CC Fragments of CD26 can be prepd in the following manner.
 CC CD26 XbaI-SphI cDNA fragment is ligated to the vector
 CC RCSR-alpha-26 XbaI-HindIII DNA fragment and the linker Q46092.
 CC The linker introduces an in-frame stop codon that results in the
 CC deletion of the segment of CD26 from AA 594 to the carboxy
 CC terminus of the wild-type protein. This deletion mutant, shown
 CC in R40916, lacks the putative catalytic site of CD26 and has a new
 CC carboxy terminus given in R40917.
 SQ Sequence 593 AA;
 SQ 28 A; 22 R; 34 N; 35 D; 0 B; 10 C; 21 Q; 31 E; 0 Z; 30 G; 11 H;
 SQ 39 I; 54 L; 34 K; 9 M; 22 F; 25 P; 49 S; 41 T; 17 W; 46 Y; 35 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:24-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || |||
RFRPSEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYISNEYKGMPG
 360      370      380      390      400      X 410      420

GRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKY
 430      440      450      460
  
```

7. US-08-249-182-10 (1-12)

R40909 Sequence encoded by human CD26 cDNA.

ID R40909 standard; Protein; 766 AA.
 AC R40909;
 DT 05-FEB-1994 (first entry)
 DE Sequence encoded by human CD26 cDNA.
 KW Human T cell activation antigen; monoclonal antibody Tai.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 7..28
 FT /label= hydrophobic
 FT Region 29..323
 FT /label= N-terminal glycosylated region of
 FT extracellular domain
 FT /note= "8 sites for N-linked glycans"
 FT Region 324..551
 FT /label= Cysteine rich region of extracellular
 FT domain
 FT /note= "1 N-linked glycosylation site"
 FT Region 552..766
 FT /label= C-terminal region of extracellular
 FT domain

FT /note= 1 N-linked glycosylation site & 1
 FT catalytic site"
 FT Active_site 627..631
 FT /label= active site of serine protease/esterase
 FT /note= "fits the consensus sequence GX SXG"
 PN W09316102-A.
 PD 19-AUG-1993.
 PF 09-APR-1992; U02892.
 PR 06-FEB-1992; US-832211.
 PA (DAND) DANA FARBER CANCER INST INC.
 PI Morimoto C, Schlossman SF, Tanaka T;
 DR WPI; 93-272827/34.
 DR N-PSDB; 046089.
 PT Polypeptide fragments of CD26 - are capable of disrupting binding
 PT of CD45 and CD26 and thus interfering with T-cell activation
 PS Disclosure; pages 39-43; 73pp; English.
 CC C26 is a human T cell activation antigen originally identified by
 CC its reactivity with the MAb Ta1. C26 cDNA library was constructed
 CC from human PHA-activated T cells using the CDM7vector. The hydrophobic
 CC N-terminal of the predicted CD26 polypeptide has the characteristics
 CC of a signal sequence of the type II membrane protein, which is
 CC reinforced by the observation that potential N-glycosylation sites
 CC are located in the carboxy side of the hydrophobic core. Therefore
 CC the N-terminal 6 AAs are predicted to be cytoplasmic, the next 22
 CC AAs are predicted to transverse the cytoplasmic membrane, and the
 CC 738 C-terminal AAs constitute the predicted extracellular domain.
 SQ Sequence 766 AA;
 SQ 40 A; 30 R; 40 N; 46 D; 0 B; 12 C; 30 Q; 40 E; 0 Z; 43 G; 19 H;
 SQ 49 I; 62 L; 40 K; 15 M; 31 F; 29 P; 64 S; 50 T; 21 W; 56 Y; 49 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:24-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || |||
 RFRPSEPHFTLDGNSFYKIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEALTSDYLYYISNEYKGMPC
 360 370 380 390 400 X 410 420
 GRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKY
 430 440 450 460

8. US-08-249-182-10 (1-12)

P50287 Sequence encoded by hepatitis A virus (HAV) cDNA f

ID P50287 standard; Protein; 854 AA.
 AC P50287;
 DT 30-NOV-1991 (first entry)
 DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the
 DE genome 5' terminus to the end of the area corresponding to the
 DE capsid protein region of poliovirus RNA.
 KW Hepatitis A virus assay; antigen; antibody.
 OS Hepatitis A virus.
 PN W08501517-A.
 PD 11-APR-1985.
 PF 27-SEP-1984; U01552.
 PR 30-SEP-1983; US-537911.
 PA (MASI) MASSACHUSETTS INST TECH.
 PI Ticehurst JR, Baltimore D, Feinstone SM, Purcell RH,
 PI Racaniello VR;
 DR WPI; 85-098846/16.
 DR N-PSDB; N50330.
 PT New hepatitis A virus CDNA - useful in assays for the virus and

PS Example; Fig 7; 60pp; English.
 CC The inventors claim HAV cDNA and a method for producing it, whereby
 CC large amts. can be obtd. economically. The cDNA is useful in the
 CC assay for detection of HAV quickly and easily and with high
 CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
 CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.
 SQ Sequence 854 AA;
 SQ 50 A; 37 R; 36 N; 43 D; 0 B; 12 C; 38 Q; 45 E; 0 Z; 53 G; 22 H;
 SQ 46 I; 70 L; 39 K; 21 M; 47 F; 45 P; 73 S; 70 T; 12 W; 34 Y; 61 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:41-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || | ||
 WRGDLVDFDQVFPTKYHSGRLLFCFVPGNELIDVSGITLKGATTAPCAVMDITGVQSTLRFVRVPWISDTPYR
 360 370 380 390 400 X 410 420
 VNRYTKSAHQKGEYTAIGKLIVYCYNRLTSPSNVASHVRV
 430 440 450 460

9. US-08-249-182-10 (1-12)

P50231 Sequence encoded by partial sequence of hepatitis

ID P50231 standard; Protein; 993 AA.
 AC P50231;
 DT 28-NOV-1991 (first entry)
 DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 OS Hepatitis A virus.
 FH Key Location/Qualifiers
 FT Protein 628..993
 FT /note= "claimed; X denotes translated stop codons
 FT and unspecified triplets"
 PN EP-138704-A.
 PD 24-APR-1985.
 PF 09-OCT-1984; 402025.
 PR 14-OCT-1983; US-541836.
 PR 02-MAR-1984; US-585942.
 PA (MERI) MERCK & CO INC.
 PI Hughes JV, Scolnick EM, Tomassini JE;
 DR WPI; 85-100818/17.
 DR N-PSDB; N50274.
 PT New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus
 PS Disclosure; Page 17-23; 49pp; English.
 CC VP1 is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 SQ Sequence 993 AA;
 SQ 58 A; 48 R; 38 N; 48 D; 0 B; 16 C; 43 Q; 47 E; 0 Z; 63 G; 26 H;
 SQ 53 I; 94 L; 43 K; 24 M; 50 F; 51 P; 81 S; 77 T; 14 W; 35 Y; 68 V;
 SQ 16 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:41-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

DIEHLTSLDFFR
 || | ||

WRGDLVDFQVFPKYHSGRLLFCFVPGNELIDVTGITLKQATTAPCAVMDITGVQSTLRFVRVPWISDTPYR
 500 510 520 530 540 X 550 X 560

VNRYTKSAHQKGEYTAIGKLIVYCYNRLTSPSNVASHVRV
 570 580 590 600

10. US-08-249-182-10 (1-12)

P98500 Partial sequence encoded by murine muscular dystro

ID P98500 standard; Protein; 1091 AA.
 AC P98500;
 DT 10-MAR-1993 (revised entry)
 DE Partial sequence encoded by murine muscular dystrophy (MD) cDNA.
 KW Dystrophin; muscular dystrophy; probe; antibody; diagnosis;
 KW prenatal; heterozygote; gene therapy; genetic screening;
 KW foetal screening.
 OS Mus musculus.
 PN W08906286-A.
 PD 13-JUL-1989.
 PF 16-DEC-1988; U04504.
 PR 22-DEC-1987; US-136618.
 PA (CHIL-) CHILDRENS MED CENT.
 PI Kunkel LM, Monaco A, Hoffman EP, Koenig M;
 DR WPI; 89-220587/30.
 DR N-PSDB; N97129.
 PT Muscular dystrophy gene - used for prepn. of probes, dystrophin
 PT polypeptide and antibodies for diagnosis and therapy of muscular
 PT dystrophy
 PS Example; Fig 7B; 68pp; English.
 CC The inventors claim an MD probe comprising a purified ss NA SQ which
 CC hybridises to at least a part of the MD gene; pure dystrophin (DS)
 CC polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The
 CC probes are equal to or greater than 10b of one of 12 cDNA sequences
 CC deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd
 CC gene.
 SQ Sequence 1091 AA;
 SQ 63 A; 42 R; 43 N; 44 D; 0 B; 13 C; 93 Q; 138E; 0 Z; 30 G; 25 H;
 SQ 44 I; 130L; 89 K; 32 M; 27 F; 27 P; 79 S; 72 T; 23 W; 18 Y; 59 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:10-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || | ||

QNGFNYSLDTVKEMAKKAPSEICQKYLSEFEEIEGHWKKLSSQLVESCQKLEEHMKNLRKFRIHIKTLQKWM
 700 710 720 730 740 X 750 760

AEVDVFLKEEWPALGDAEILKKQLKQCRLLVGDITQIQPS
 770 780 790 800

11. US-08-249-182-10 (1-12)

R12141 Enteroviral polypeptide.

ID R12141 standard; Protein; 2185 AA.
 AC R12141;
 DT 05-AUG-1991 (first entry)
 DE Enteroviral polypeptide.
 KW Enteroviruses; monoclonal antibodies; myocarditis; myositis;

KW meningitis; encephalitis; pancreatitis; post viral fatigue.
 OS Enterovirus sp.
 PN DE3939200-A.
 PD 29-MAY-1991.
 PF 27-NOV-1989; 939200.
 PR 27-NOV-1989; DE-939200.
 PA (PLAC) MAX PLANCK GES WISSENSCH.
 PI Kandolf R;
 DR WPI; 91-165150/23.
 DR N-PSDB; Q11816.
 PT New enteroviral polypeptide for raising group specific antibodies
 PT - for detecting any type of enterovirus in blood or serum, and
 PT new DNA encoding it
 PS Claim 1; pages 14-15; 26pp; German.
 CC This eteroviral polypeptide is used to raise poly- or monoclonal
 CC antibodies (Abs). These are useful in assays for detecting entero-
 CC virus specific antigens, as an indication of enteroviral disease.
 CC All 70 serotypes of the enteroviral family can be detected.
 CC Diseases such as myocarditis, myositis, meningitis, encephalitis
 CC and pancreatitis can be diagnosed using the Abs.
 SQ Sequence 2185 AA;
 SQ 148A; 89 R; 107N; 100D; 0 B; 52 C; 92 Q; 117E; 0 Z; 155G; 49 H;
 SQ 115I; 174L; 118K; 61 M; 96 F; 120P; 154S; 147T; 32 W; 93 Y; 166V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:20-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 41% Matches = 5 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 ||| ||
 LMVIPFVPLDYCPGSTTYPITVTIAPMCAEYNGRLRAGHGLPTMNTPGSCQFLTSDDFQSPSAMPQYDVT
 300 310 320 330 340 X 350 X 360
 PEMRIPGEVKNLMEIAEVDVVVPQNVGEKVNSEAYQIP
 370 380 390 400

12. US-08-249-182-10 (1-12)

P30304 Sequence corresp. to a protein antigen or allergen

ID P30304 standard; Protein; 12 AA.
 AC P30304;
 DT 20-APR-1992 (first entry)
 DE Sequence corresp. to a protein antigen or allergen with high local
 DE average hydrophilicity
 KW Synthetic vaccine; antigen; allergen; immunological response;
 KW antibody.
 PN EP--93851-A.
 PD 16-NOV-1983.
 PF 11-MAR-1983; 102392.
 PR 09-JAN-1981; US-223558.
 PR 12-JUN-1981; US-272855.
 PR 15-MAR-1982; US-358150.
 PR 28-JAN-1983; US-461802.
 PR 16-DEC-1986; US-942562.
 PA (NYBL-) NEW YORK BLOOD CENT.
 PI Hopp TP;
 DR WPI; 83-822049/47.
 PT Synthetic vaccine - contains peptide residue coupled to higher
 PT alkyl or alkenyl Gps. and with 6 amino acids in residue
 PS Claim 59; Page 48; 54pp; English.
 CC The inventors claim a synthetic vaccine which comprises a peptide
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C, or
 CC other lipophilic substance. The residue contains a sequence of 6 AAs

CC corresp.to the SQ of such AAs in a protein antigen or allergen where
 CC the greatest local average hydrophilicity is found. Pref. the AAs in
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is
 CC pref. coupled to the terminal amino gp. of the residue opt. via a CO
 CC gp. If a lipophilic substance is used, it is pref. palmitic, stearic,
 CC behenic, oleic or mycolic acid.
 SQ Sequence 12 AA;
 SQ 0 A; 1 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 0 G; 2 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 0 P; 1 S; 0 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:01-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.47
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X 10
 DIEHLTSLDFFR
 ||||
 HDFSDEIEHLV
 X 10 X

13. US-08-249-182-10 (1-12)

R41129 HCV peptide XXc-1 (aa 383-404; E2/NS1 N-terminal).

ID R41129 standard; peptide; 24 AA.
 AC R41129;
 DT 22-MAR-1994 (first entry)
 DE HCV peptide XXc-1 (aa 383-404; E2/NS1 N-terminal).
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;
 FT A (optional)= one or more amino acids, NH2 or
 FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH2, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 24
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN W09318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT

Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      ||      ||
SQFGQILDILVSRSLKMRGQAFVI
      10      20

```

15. US-08-249-182-10 (1-12)
R25061 Thyrotrope reacting peptide 1.

ID R25061 standard; peptide; 24 AA.
AC R25061;
DT 11-DEC-1992 (first entry)
DE Thyrotrope reacting peptide 1.
KW autoantibody; thyroid functional aberration; hypothyroidism.
OS Synthetic.
PN CA2050046-A.
PD 13-APR-1992.
PF 27-AUG-1991; 050046.
PR 12-OCT-1990; JP-274668.
PA (TANA) TANABE SEIYAKU CO.
PI Mori M, Murakami M;
DR WPI; 92-217472/27.
PT New peptide(s) which react with thyrotrope antibody - used for
PT the diagnosis of diseases of thyroid functional aberration such
PT as hyperthyroidism
PS Claim 1; Page 15; 19pp; English.
CC The sequences give in R25061-2 are novel peptides which react
CC immunochemically with an antibody of thyrotrope. The antibodies
CC raised against these peptides and other sequences containing these
CC peptides can be used to isolate autoantibodies which lead to
CC diseases of thyroid functional aberration, such as neonatal
CC hypothyroidism or hyperthyroidism.
SQ Sequence 24 AA;
SQ 0 A; 2 R; 0 N; 2 D; 0 B; 1 C; 3 Q; 2 E; 0 Z; 0 G; 1 H;
SQ 2 I; 1 L; 1 K; 0 M; 1 F; 3 P; 2 S; 2 T; 0 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:14-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.47
Residue Identity = 33% Matches = 4 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      ||      ||
HQEEDFRVTCKDIQRIPSLPPSTQ
      10 X      20 X

```

> 0 <
0| 10 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-08-249-182-11.res made by on Wed 21 Sep 94 12:07:16-PDT.

Query sequence being compared:US-08-249-182-11 (1-23)
Number of sequences searched: 5543
Number of scores above cutoff: 3892

Results of the initial comparison of US-08-249-182-11 (1-23) with:
File : /home/shears/loring/lorin*.pep

PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are
 CC useful for the detection of antibodies to HCV, and/or HIV,
 CC and/or HTLV-I or II.
 SQ Sequence 24 AA;
 SQ 0 A; 2 R; 0 N; 0 D; 0 B; 1 C; 1 Q; 0 E; 0 Z; 4 G; 2 H;
 SQ 0 I; 2 L; 0 K; 0 M; 1 F; 0 P; 1 S; 5 T; 0 W; 0 Y; 3 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:31-PDT using FindSeq

Initial Score = 4 Optimized Score = 5 Significance = 2.47
 Residue Identity = 50% Matches = 6 Mismatches = 4
 Gaps = 2 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
          |||| ||
XGHTRVTGGVQGHVTC TLTSL--FRX
      10 X      20 X
  
```

14. US-08-249-182-10 (1-12)

R39589 Ribonucleoprotein snRNP-U1 protein A peptide.

ID R39589 standard; peptide; 24 AA.
 AC R39589;
 DT 15-NOV-1993 (first entry)
 DE Ribonucleoprotein snRNP-U1 protein A peptide.
 KW In vitro diagnosis; mixed connective tissue disease; MCTD;
 KW autoantibodies kit; complex formation; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Peptide 4..16
 FT /note= "preferred peptide"
 PN FR2682113-A.
 PD 09-APR-1993.
 PF 25-JUN-1991; 007818.
 PR 25-JUN-1991; FR-007818.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PI Barakat S, Briand JP, Muller S, van Regenmortel MHV.
 DR WPI; 93-215860/27.
 PT Peptide(s) derived from ribonucleoprotein snRNP-U1 - for
 PT diagnosing auto:immune diseases and treating e.g. connective
 PT tissue disorders
 PS Claim 1; Page 30; 37pp; French.
 CC The peptide is derived from ribonucleoprotein snRNP-U1 and is useful
 CC for in vitro diagnosis of mixed connective tissue disease (MCTD) by
 CC formation of a complex with autoantibodies (AAb). It is specific for
 CC AAb present in patients with MCTD and is not significantly homologous
 CC with other antigens associated with autoimmune disease. It may also
 CC be used in vaccines (usual dose 0.01-100mg/kg) while antibodies are
 CC useful as specific immunological probes for rMCTD and as agents for
 CC raising anti-idiotypic antibodies which are also useful in diagnosis
 CC and vaccines.
 SQ Sequence 24 AA;
 SQ 1 A; 2 R; 0 N; 1 D; 0 B; 0 C; 3 Q; 0 E; 0 Z; 2 G; 0 H;
 SQ 3 I; 3 L; 1 K; 1 M; 2 F; 0 P; 3 S; 0 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.47
 Residue Identity = 33% Matches = 4 Mismatches = 8

Number of residues: 482838
 Number of sequences searched: 5543
 Number of scores above cutoff: 3892

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

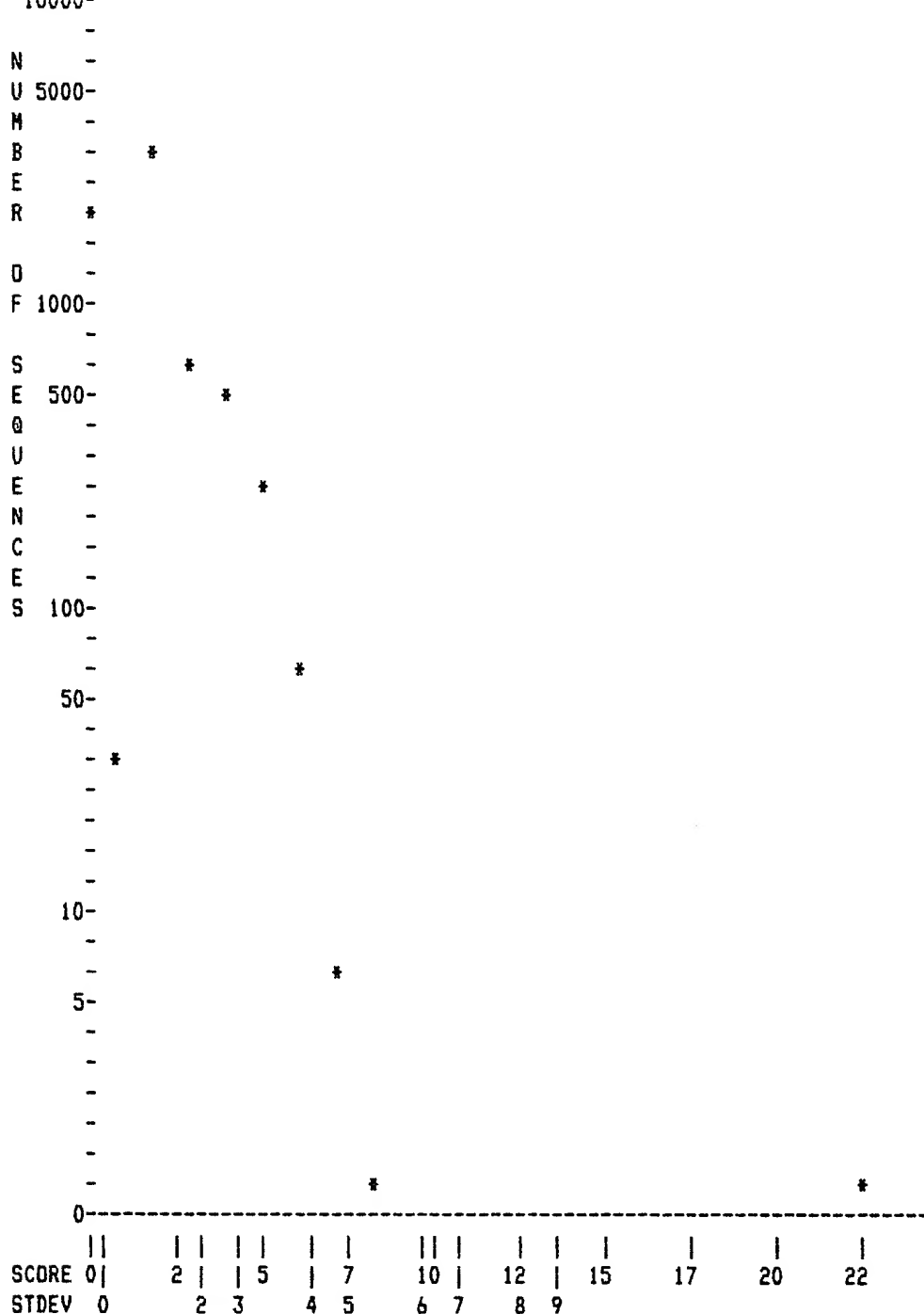
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 14 standard deviations above mean ****						
1. R37453	Autotaxin peptide ATX 103.	23	22	22	14.17	0
**** 4 standard deviations above mean ****						
2. R04574	Derived amino acid sequence o	810	8	9	4.72	0
3. P71672	Human serine protease.	262	7	7	4.05	0
4. R33279	43 kD endoflagellum sheath pr	320	7	9	4.05	0
5. R07130	H20B receptor.	392	7	7	4.05	0
6. R07131	H20A receptor.	416	7	7	4.05	0
7. R31041	srnR polypeptide.	655	7	7	4.05	0
8. FRZE_MYXXA	GLIDING MOTILITY REGULATORY P	777	7	7	4.05	0
**** 3 standard deviations above mean ****						
9. R27328	Peptide corresp. to an epitop	19	6	7	3.37	0
10. R37537	Methicillin-resistant S. aure	32	6	7	3.37	0
11. R37524	Methicillin-resistant S. aure	32	6	7	3.37	0
12. R27562	Insert B to prevent steric hi	34	6	7	3.37	0
13. R40101	Hib OMP P1-P2 hybrid peptide	54	6	7	3.37	0
14. R43686	Human kappa constant domain a	106	6	6	3.37	0
15. R41687	Undefined ORF1 encoded by pla	106	6	6	3.37	0
16. P81028	C region of L chain (chi) of	106	6	6	3.37	0
17. R11013	Feline immunoglobulin kappa c	109	6	6	3.37	0
18. R39464	MAB BW 2128 heavy chain V-reg	119	6	6	3.37	0
19. R38883	Antibody light chain.	129	6	6	3.37	0
20. R31531	H1223 MAB light chain.	130	6	6	3.37	0
21. P70624	Sequence encoded by anti-hepa	136	6	6	3.37	0
22. R37717	Mouse 4C10 anti-idiotypic Ab h	138	6	6	3.37	0
23. R33953	gH1 variable domain.	139	6	6	3.37	0
24. R27049	VH425 antibody cloned into pU	140	6	7	3.37	0
25. R22564	Antibody specific for PB1.	178	6	6	3.37	0
26. R05274	Segment of human B cell diffe	185	6	7	3.37	0
27. R05275	Segment of human B cell diffe	185	6	7	3.37	0
28. P90436	Interferon-beta-2.	212	6	7	3.37	0
29. R43338	Completely humanised C4G1 Ig	214	6	6	3.37	0
30. R30776	H52L6-158 murine anti-CD18 an	214	6	6	3.37	0
31. R33312	Humanised MaE11 Version 1 (in	218	6	6	3.37	0
32. R04494	HIV fusion protein PB1rf	229	6	7	3.37	0
33. R04496	HIV fusion protein PB1sc.	231	6	7	3.37	0
34. R30777	pH52-9.0 humanised murine ant	233	6	6	3.37	0
35. R22755	Reshaped CD4 antibody light c	233	6	6	3.37	0
36. R22754	Reshaped CAMPATH-1 antibody l	233	6	6	3.37	0
37. R20058	Light chain of 3D6 anti-HIV a	234	6	6	3.37	0
38. R13050	CD4-specific CDR-grafted ligh	234	6	6	3.37	0
39. R42065	Human anti-HBs light chain.	236	6	6	3.37	0
40. R22565	V1lys-HuCKappa region of Fab	236	6	6	3.37	0

1. US-08-249-182-11 (1-23)

R37453 Autotaxin peptide ATX 103.

ID R37453 standard; peptide; 23 AA.

AC R37453;



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.48

Times:	CPU	Total Elapsed
	00:00:08.93	00:00:09.00

DE Autotaxin peptide ATX 103.
 KW Cell motility stimulating; cancer metastasis; antibody; detection;
 KW immunostains; disease outcome prediction; therapy choice;
 KW cancer therapy; crosslinked toxins.
 OS Synthetic.
 PN US7822043-A.
 PD 01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 103. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 23 AA;
 SQ 0 A; 1 R; 2 N; 2 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 2 G; 0 H;
 SQ 1 I; 4 L; 0 K; 0 M; 1 F; 1 P; 1 S; 4 T; 0 W; 1 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 22 Optimized Score = 22 Significance = 14.17
 Residue Identity = 95% Matches = 22 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

X      10      20 X
TEFLSNYLTNVDDITLVPETLGR
|||||
TEFLSNYLTNVDDITLVPGLGR
X      10      20 X

```

2. US-08-249-182-11 (1-23)

R04574 Derived amino acid sequence of coding region of mu

ID R04574 standard; protein; 810 AA.
 AC R04574;
 DT 17-SEP-1990 (first entry)
 DE Derived amino acid sequence of coding region of murine IL-4 receptor
 KW mammalian interleukin-4 receptor; cytokine; antibody production;
 OS synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 209..232
 FT /label= putative transmembrane region
 PN EP-367566-A.
 PD 09-MAY-1990.
 PF 31-OCT-1989; 311244.
 PR 23-JUN-1989; US-370924.
 PA (IMMU-) Immunex Corp.
 PI Cosman DJ, Park L, Mosley B, Beckmann P, March CJ, Idzerda R,
 DR WPI; 90-141470/19.
 DR N-PSDB; 004305.
 PT Recombinant mammalian interleukin-4 receptor used in diagnosis,
 PT assays and therapy and for prodn. of antibodies for diagnosis,therapy
 PT and for prodn. of antibodies
 PS Disclosure; p; English.
 CC The interleukin-4 receptor can be used to regulate immune responses or to
 CC treat IgE-induced hypersensitivity.

CC see also 604307.
SQ Sequence 810 AA;
SQ 49 A; 23 R; 28 N; 34 D; 0 B; 32 C; 39 Q; 52 E; 0 Z; 63 G; 18 H;
SQ 29 I; 80 L; 30 K; 15 M; 25 F; 83 P; 89 S; 39 T; 16 W; 18 Y; 48 V;
CC Retrieved by shears on Wed 21 Sep 94 11:56:58-PDT using FindSeq

Initial Score = 8 Optimized Score = 9 Significance = 4.72
Residue Identity = 43% Matches = 10 Mismatches = 12
Gaps = 1 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
|| || | | || ||
AFSSLLSSNAIRGDTAAAGTDDGHGGYKPFQNPVPNGSPSSVPLFTFGLDTE-LSPSPNSDPKSPPECLG
610 620 630 640 650 X 660 670

X
R

LELGLKGGDWVKAPPPADEVPKPFDDLGFGIVYSSLTCHLCGHLKQHSQ
X 680 690 700 710 720

3. US-08-249-182-11 (1-23)

P71672 Human serine protease.

ID P71672 standard; Protein; 262 AA.
AC P71672;
DT 10-JUN-1991 (first entry)
DE Human serine protease.
KW Serine protease; assay; antibodies; immunisation; HSP; diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..28
FT /label= sig_peptide
FT Protein 28
FT /label= mat_protein
PN EP-245051-A.
PD 11-NOV-1987.
PF 01-MAY-1987; 303945.
PR 06-MAY-1986; US-860085.
PR 08-MAY-1986; US-861221.
PR 31-DEC-1986; US-948248.
PA (STRD) Leland Stanford JR Univ.
PI Weissman IL, Gershenfeld HK.
DR WPI; 87-315213/45.
DR N-PSDB; N71407.
PT New pure human serine protease and fragments - used as labels in
PT assays and for prodn. of antibodies for passive immunisation
PT against immune disorders.
PS Disclosure; Fig 1; 7pp; English.
CC Amino acid homology within the active enzyme to the mouse protease is
CC 71% with 77% at the DNA level. The overall homolgy is 72% when the
CC complete coding region and the 3' UTR are included. The amino acids
CC of the charge-relay system are His41, Asp86 and Ser184. The acidic
CC residue Asp178 determines substrate specificity for Lys or Arg.
CC Asn142 is a Asn-linked carbohydrate site.
CC The HSP is produced by activated killer cells. The enzyme acts in
CC conjunction with other components of a killer cell to provide
CC cytolytic capability.
SQ Sequence 262 AA;
SQ 13 A; 15 R; 15 N; 13 D; 0 B; 11 C; 4 Q; 11 E; 0 Z; 22 G; 7 H;
SQ 18 I; 27 L; 19 K; 7 M; 5 F; 13 P; 18 S; 14 T; 3 W; 7 Y; 20 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:14-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 4.05

Residue Identity = 30% Matches = 7 Mismatches = 18
Gaps = 0 Conservative Substitutions = 0

```

      X      10      20  X
      TEFLSNYLTNVDDITLVPETLGR
      ||  |  |  ||
MRNSYRFLASSLSVVVSLLLIPEDVCEKIIGGNEVTPHSRPMVLLSLDRKTICAGALIAKDWLTAHCNL
      X      10      20      X 30      40      50      60      70
NKRSQ
```

4. US-08-249-182-11 (1-23)

R33279 43 kD endoflagellum sheath protein.

ID R33279 standard; Protein; 320 AA.
AC R33279;
DT 16-JUL-1993 (first entry)
DE 43 kD endoflagellum sheath protein.
KW Endoflagellum; sheath protein; T. hyodysenteriae; core; antibody;
KW bacteriacide; 43 kD; vaccine; infection; swine dysentery.
OS Treponema hyodysenteriae.
FH Key Location/Qualifiers
FT Peptide 1..19
FT /note= "Signal peptide"
FT Protein 20..320
FT /note= "Mature protein"
PN EP-534526-A.
PD 31-MAR-1993.
PF 14-SEP-1992; 202796.
PR 25-SEP-1991; EP-202478.
PR 24-JUL-1992; EP-202273.
PA (DUIN) DUPHAR INT RES BV.
PI Koopman MBH, Kusters JG;
DR WPI; 93-102665/13.
DR N-PSDB; 038583.
PT Vaccine to protect pigs against swine dysentery - comprises
PT Treponema hyodysenteriae endo-flagellum sheath protein, applied
PT orally or intranasally
PS Claim 2; Page 21-22; 34pp; English.
CC This sequence represents the endoflagellum sheath protein of T.
CC hyodysenteriae. The endoflagellum consists of at least four
CC proteins, this protien forms the sheath of the flagellum and three
CC proteins, of molecular weights 37, 34 and 32 kD, make up its core.
CC Antibodies raised against the sheath protein have been shown to be
CC bacteriacidal for T. hyodysenteriae. The 43 kD sheath protein can
CC be used in the production of a vaccine against infections such as
CC swine dysentery.
SQ Sequence 320 AA;
SQ 28 A; 18 R; 23 N; 22 D; 0 B; 0 C; 11 Q; 25 E; 0 Z; 26 G; 1 H;
SQ 13 I; 26 L; 15 K; 4 M; 10 F; 7 P; 16 S; 20 T; 8 W; 17 Y; 30 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:52-PDT using FindSeq

Initial Score = 7 Optimized Score = 9 Significance = 4.05
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

```

      X      10      20
      TEFLSNYLTNVDDITLVPETLG
      || || | || |
IDNVGEIKSISSWVYGRNYLISYFVNLGNEFGELKSYPMGTVVYFNGWRQVRWENREYLPNVRDSVLVREPLY
      140      150      160      170      180      190      200
```

X
R

PRMIPSVKLDLGLFYRTKDTKGGDFITYVKDVTLEYDVVVVDFFEDIDDEA
210 220 230 240 250

5. US-08-249-182-11 (1-23)

R07130 H20B receptor.

ID R07130 standard; protein; 392 AA.
AC R07130;
DT 23-JAN-1991 (first entry)
DE H20B receptor.
KW Picornavirus proteins; poliovirus; transgenic animals; vaccines;
KW antibodies; imaging.
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label=signal peptide
FT Domain 345..368
FT /label=transmembrane domain
PN W09010699-A.
PD 20-SEP-1990.
PF 09-MAR-1990; U01320.
PR 10-MAR-1989; US-321957.
PA (UYCD-) COLUMBIA UNIV NY.
PI Racaniello V, Mendelsohn C, Costantini F;
DR WPI; 90-305023/40.
DR N-PSDB; Q06069.
PT DNA encoding picornavirus partic. poliovirus receptors proteins -
PT for treating picornavirus infections or for expression in
PT transgenic animals used to test vaccines
PS Disclosure; fig 4; 88pp; English.
CC This poliovirus receptor, H20B, has a sequence differing from the
CC receptor, H20A at the cytoplasmic tail only. Antibodies (Abs)
CC raised against it are useful for targetted delivery of the human
CC poliovirus, conjugated to a drug. Transgenic animals contg. the
CC corresp. DNA (genomic- or cDNA) can be used to test the efficiency
CC and virulence of picornavirus vaccines.
CC See also Q06070.
SQ Sequence 392 AA;
SQ 26 A; 17 R; 17 N; 9 D; 0 B; 10 C; 20 Q; 20 E; 0 Z; 30 G; 10 H;
SQ 13 I; 42 L; 9 K; 9 M; 12 F; 31 P; 31 S; 28 T; 10 W; 10 Y; 38 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:08-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 4.05
Residue Identity = 30% Matches = 7 Mismatches = 16
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
||| | |||
AKPQNTAEVQKVQLTGEPVPMARCVSTGGRPPAQITWHSDLGGMPNTSQVPGFLSGTVTVTSLWILVPSSQV
150 160 170 180 190 X 200 210

X
R

DGKNVTCKVEHESFEKPQLLTVNLTVYYPPEVSISGYDNNWYLGQNEATLT
X 220 230 240 250 260

6. US-08-249-182-11 (1-23)

R07131 H20A receptor.

ID R07131 standard; protein; 416 AA.
AC R07131;
DT 23-JAN-1991 (first entry)
DE H20A receptor.

KW Picornavirus proteins; poliovirus; transgenic animals; vaccines;
 KW antibodies; imaging.
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label=signal peptide
 FT Domain 345..368
 FT /label=transmembrane domain
 PN W09010699-A.
 PD 20-SEP-1990.
 PF 09-MAR-1990; U01320.
 PR 10-MAR-1989; US-321957.
 PA (UYCO-) COLUMBIA UNIV NY.
 PI Racaniello V, Mendelsohn C, Costantini F;
 DR WPI; 90-305023/40.
 DR N-PSDB; Q06070.
 PT DNA encoding picornavirus partic. poliovirus receptors proteins -
 PT for treating picornavirus infections or for expression in
 PT transgenic animals used to test vaccines
 PS Disclosure; fig 4; 88pp; English.
 CC This poliovirus receptor, H20A, has a sequence differing from the
 CC receptor, H20B at the cytoplasmic tail only. Antibodies (Abs)
 CC raised against it are useful for targetted delivery of the human
 CC poliovirus, conjugated to a drug. Transgenic animals contg. the
 CC corresp. DNA (genomic- or cDNA) can be used to test the efficiency
 CC and virulence of picornavirus vaccines.
 CC See also Q06069.
 SQ Sequence 416 AA;
 SQ 30 A; 18 R; 18 N; 10 D; 0 B; 9 C; 21 Q; 22 E; 0 Z; 32 G; 10 H;
 SQ 13 I; 42 L; 9 K; 9 M; 12 F; 32 P; 37 S; 31 T; 10 W; 11 Y; 40 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:08-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 4.05
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTVDDITLVPETLG
 ||| | |||
 AKPQNTAEVQKVQLTGEPVPMARCVSTGGRPPAQITWHSDLGMPNTSQVPGFLSGTVTVTSLWILVPSSQV
 150 160 170 180 190 X 200 210

X
 R

DGKNVTCKVEHESFEKPQLLTVNLTVVYPPEVSISGYDNNWYLGQNEATLT
 X 220 230 240 250 260

7. US-08-249-182-11 (1-23)

R31041 srnR polypeptide.

ID R31041 standard; Protein; 655 AA.
 AC R31041;
 DT 24-MAY-1993 (first entry)
 DE srnR polypeptide.
 KW Regulatory; activator; protein; srnR; complementation experiment;
 KW macrolide; biosynthetic; gene; transcription; S. ambofaciens;
 KW biosynthesis; antibody; polyclonal; monoclonal; pathway.
 OS Streptomyces ambofaciens.
 PN EP-524832-A.
 PD 27-JAN-1993.
 PF 24-JUL-1992; 306792.
 PR 26-JUL-1991; US-736178.
 PA (ELIL) LILLY & CO ELI.
 PI Rao RN, Turner JR;

DR N-PSDB; 035141.
PT DNA encoding regulatory activator protein *srnR* - for macrolide
PT biosynthesis to increase efficiency of antibiotic prodn.
PS Claim 5; Page 19-20; 22pp; English.
CC The sequence given represents regulatory (activator) protein *srnR*.
CC The DNA encoding this peptide has the ability in complementation
CC experiments to restore macrolide biosynthetic gene transcription
CC in mutants having defective *srnR* genes due to insertional
CC inactivation of that region of the *S. ambofaciens* genome. The *srnR*
CC gene product activates macrolide biosynthetic gene transcription and
CC may be used to increase the efficiency of macrolide biosynthesis.
CC The translation product of this gene is useful for the generation of
CC antibodies (polyclonal or monoclonal) which are useful in the
CC detection of other macrolide biosynthetic pathways.
SQ Sequence 655 AA;
SQ 84 A; 69 R; 8 N; 48 D; 0 B; 8 C; 14 Q; 45 E; 0 Z; 56 G; 16 H;
SQ 13 I; 86 L; 6 K; 10 M; 13 F; 37 P; 38 S; 37 T; 3 W; 12 Y; 52 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:38-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 4.05
Residue Identity = 30% Matches = 7 Mismatches = 16
Gaps = 0 Conservative Substitutions = 0

					X	10	20
					TEFLSNYL	TNVDDITL	VPETLG
PLGRPLAEYGT	LRPVAPTEL	RAACRRAAET	GRPTSVAPGV	WTVPLLP	GGNAGFLL	TDLGPDADHT	AVPLLPM
310	320	330	340	350	X 360	370	

X						
R						
VARTLALH	LRVQHD	SPKAQSHQ	EFFDDL	IGAPRSPT	LLRERAL	MFSHSFR
380	390	400	410	420		

8. US-08-249-182-11 (1-23)
FRZE_MYXXA GLIDING MOTILITY REGULATORY PROTEIN (EC 2.7.1.-).

ID FRZE_MYXXA STANDARD; PRT; 777 AA.
AC P18769;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)
DE GLIDING MOTILITY REGULATORY PROTEIN (EC 2.7.1.-).
GN FRZE.
OS MYXOCOCCUS XANTHUS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; MYXOBACTERIALES;
OC MYXOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 90332690
RA MCCLEARY W.R., ZUSMAN D.R.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:5898-5902(1990).
RN [2]
RP PHOSPHORYLATION AT ASN-49.
RM 91072208
RA MCCLEARY W.R., ZUSMAN D.R.;
RL J. BACTERIOL. 172:6661-6668(1990).
CC -!- FUNCTION: FRZE IS INVOLVED IN A SENSORY TRANSDUCTION PATHWAY THAT
CC CONTROLS THE FREQUENCY AT WHICH CELLS REVERSE THEIR GLIDING
CC DIRECTION. FRZE SEEMS TO BE CAPABLE OF AUTOPHOSPHORYLATING ITSELF
CC ON AN HISTIDINE RESIDUE AND THEN TO TRANSFER THAT GROUP TO AN
CC ASPARTATE RESIDUE IN THE C-TERMINAL PART OF THE PROTEIN.
CC -!- SIMILARITY: TO OTHER PROKARYOTIC REGULATORY PROTEINS WHICH BELONG

CC TO A TWO-COMPONENT REGULATORY SYSTEM AND TRANSDUCE ENVIRONMENTAL
 CC SIGNALS TO TRANSCRIPTIONAL APPARATUS.
 CC -!- SIMILARITY: FRZE IS SIMILAR TO BOTH CHEA AND CHEY.
 DR EMBL; M35192; MXFRZE.
 DR PIR; A35966; A35966.
 KW SENSORY TRANSDUCTION; TRANSFERASE; KINASE; PHOSPHORYLATION.
 FT MOD_RES 49 49 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT DOMAIN 130 197 ALA/PRO-RICH (POSSIBLE HINGE REGION).
 SQ SEQUENCE 777 AA; 83189 MW; 2825541 CN;
 CC -!- Retrieved by shears on Wed 21 Sep 94 11:59:19-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 4.05
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYLTVDDITLVPETLG
                                ||      |      |||
VEDDGRGIDPERLRQAAISKRLINAVQAAALSEREAIELIFRPGFSTRDQVSELSGRGVGMDVVKRKVETLG
  420      430      440      450      460      X 470      480

X
R

GSVGVSSRIGRGSTITRLRP@SLALMKVLLVRLGDDVYGMPAADVEAVMRV
  490      500      510      520      530

```

9. US-08-249-182-11 (1-23)

R27328 Peptide corresp. to an epitope of HIV-1 gp120 prot

ID R27328 standard; peptide; 19 AA.
 AC R27328;
 DT 27-OCT-1992 (first entry)
 DE Peptide corresp. to an epitope of HIV-1 gp120 protein.
 KW Human immunodeficiency virus; vaccine; env gene; HTLV-IIIB;
 KW immunisation; AIDS; antibodies.
 OS Human immunodeficiency virus.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "if presenet is an amino acid selected to
 FT facilitate coupling of the peptide to a
 FT carrier."
 FT Misc_difference 19
 FT /note= "may be absent"
 FT Modified_site 19
 FT /note= "may be amidated"
 PN W09205800-A.
 PD 16-APR-1992.
 PF 25-SEP-1991; SE0641.
 PR 27-SEP-1990; US-589422.
 PA (SYNT-) SYNTELLO VACCINE DE.
 PI Horal P, Jeansson S, Rymo L, Svennerholm B, Vahlne A;
 DR WPI; 92-150590/18.
 PT New peptide(s) corresp. to epitope(s) of HIV-1 gp120 protein -
 PT used in vaccination and induction of neutralising antibodies
 PT against HIV
 PS Disclosure; Page 16; 48pp; English.
 CC The peptide gp120-32 corresponds to an epitope of the gp 120
 CC protein from amino acids 400-417 encoded by the env gene of HIV-1.
 CC The peptide may be synthesised by standard solid phase techniques
 CC and is useful in a vaccine to immunise against HIV infection.
 CC Polyclonal and monoclonal antibodies are elicited by the peptide,
 CC which can thus heighten immune response in subjects already infected
 CC by HIV. The peptide may be covalently attached to similar peptides
 CC or to a carrier for use in vaccines.

CC See also R25070-3 and R27301-38.

SQ Sequence 19 AA;

SQ 0 A; 0 R; 2 N; 1 D; 0 B; 1 C; 0 Q; 2 E; 0 Z; 2 G; 0 H;

SQ 1 I; 1 L; 0 K; 0 M; 0 F; 1 P; 3 S; 4 T; 0 W; 0 Y; 0 V;

SQ 1 Others;

CC Retrieved by shears on Wed 21 Sep 94 11:58:04-PDT using FindSeq

Initial Score = 6 Optimized Score = 7 Significance = 3.37

Residue Identity = 44% Matches = 8 Mismatches = 9

Gaps = 1 Conservative Substitutions = 0

```

      X      10      20
      TEFLSNYLTVDDITLVPETLGR
      ||  ||      |  |||
XSTE-GSNNTEGSDTITLPC
      X      10      X
```

10. US-08-249-182-11 (1-23)

R37537 Methicillin-resistant *S. aureus* detection peptide.

ID R37537 standard; peptide; 32 AA.

AC R37537;

DT 08-SEP-1993 (first entry)

DE Methicillin-resistant *S. aureus* detection peptide.

KW MRSA; anti-MRSA antibody; competitive reaction; dermatitis;

KW pneumonia; organ failure; treatment; prophylaxis.

OS Synthetic.

PN DE4238806-A.

PD 27-MAY-1993.

PF 17-NOV-1992; 238806.

PR 18-NOV-1991; JP-354087.

PR 10-APR-1992; JP-134097.

PA (DAIN-) DAINABOT CO LTD.

PI Nagata M, Saito M, Sekiguchi K, Yajima R.

DR WPI; 93-176824/22.

PT Methicillin-resistant *Staphylococcus aureus* (MRSA) detection

PT method - by detection of MRSA penicillin binding protein (PBP)2

PT antigen-antibody complex, and synthetic MRSA PBP2 peptides

PS Claim 1; Page 24; 36pp; German.

CC The sequence is that of a peptide which may be used in a detection

CC method for methicillin resistant *Staphylococcus aureus* (MRSA) or its

CC fragments in a sample. The method comprises the competitive reaction

CC of the peptide with an anti-MRSA antibody present in the sample,

CC forming an immunological complex, and detection of the complex.

CC The synthetically produced peptide provides a simple and accurate

CC MRSA-test. *S. aureus* causes dermatitis, pneumonia and organ failure,

CC detection of MRSA (and therefore MRSA infections) is essential for

CC appropriate prophylaxis and treatment of the condition.

SQ Sequence 32 AA;

SQ 1 A; 0 R; 2 N; 1 D; 0 B; 1 C; 2 Q; 2 E; 0 Z; 1 G; 0 H;

SQ 2 I; 4 L; 5 K; 0 M; 1 F; 2 P; 2 S; 5 T; 0 W; 1 Y; 0 V;

CC Retrieved by shears on Wed 21 Sep 94 11:59:09-PDT using FindSeq

Initial Score = 6 Optimized Score = 7 Significance = 3.37

Residue Identity = 30% Matches = 7 Mismatches = 16

Gaps = 0 Conservative Substitutions = 0

```

      X      10      20      X
      TEFLSNYLTVDDITLVPETLGR
      ||      |  |  ||  |
YNKLTEDKKEPLLNKFQITSPGSTQKILTAC
      X      10      20      X 30
```

11. US-08-249-182-11 (1-23)

ID R37524 standard; peptide; 32 AA.
 AC R37524;
 DT 08-SEP-1993 (first entry)
 DE Methicillin-resistant *S. aureus* detection peptide.
 KW MRSA; anti-MRSA antibody; competitive reaction; dermatitis;
 KW pneumonia; organ failure; treatment; prophylaxis.
 OS Synthetic.
 PN DE4238806-A.
 PD 27-MAY-1993.
 PF 17-NOV-1992; 238806.
 PR 18-NOV-1991; JP-354087.
 PR 10-APR-1992; JP-134097.
 PA (DAIN-) DAINABOT CO LTD.
 PI Nagata M, Saito M, Sekiguchi K, Yajima R.
 DR WPI; 93-176824/22.
 PT Methicillin-resistant *Staphylococcus aureus* (MRSA) detection
 PT method - by detection of MRSA penicillin binding protein (PBP)2
 PT antigen-antibody complex, and synthetic MRSA PBP2 peptides
 PS Disclosure; Page 20; 36pp; German.
 CC The sequence is that of a peptide which may be used in a detection
 CC method for methicillin resistant *Staphylococcus aureus* (MRSA) or its
 CC fragments in a sample. The method comprises the competitive reaction
 CC of the peptide with an anti-MRSA antibody present in the sample,
 CC forming an immunological complex, and detection of the complex.
 CC The synthetically produced peptide provides a simple and accurate
 CC MRSA-test. *S. aureus* causes dermatitis, pneumonia and organ failure,
 CC detection of MRSA (and therefore MRSA infections) is essential for
 CC appropriate prophylaxis and treatment of the condition.
 SQ Sequence 32 AA;
 SQ 1 A; 0 R; 2 N; 1 D; 0 B; 0 C; 2 Q; 2 E; 0 Z; 1 G; 0 H;
 SQ 2 I; 4 L; 5 K; 1 M; 1 F; 2 P; 2 S; 5 T; 0 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:09-PDT using FindSeq

Initial Score = 6 Optimized Score = 7 Significance = 3.37
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

```

      X      10      20 X
      TEFLSNYLTVDDITLVPETLGR
      ||      ||      ||
YNKLTEDKKEPLLNFQITTPGSGTQKILTAM
      X      10      20      X 30
  
```

12. US-08-249-182-11 (1-23)

R27562 Insert B to prevent steric hindrance & competition

ID R27562 standard; Protein; 34 AA.
 AC R27562;
 DT 26-FEB-1993 (first entry)
 DE Insert B to prevent steric hindrance & competition with *E. coli* coat.
 KW Dicistronic expression vector; fusion PCR; antibody; cDNA library;
 KW ss.
 OS Synthetic.
 PN W09215678-A.
 PD 17-SEP-1992.
 PF 27-FEB-1992; U01475.
 PR 01-MAR-1991; US-663442.
 PA (STRA-) STRATAGENE.
 PI Sorge JA;
 DR WPI; 92-331724/40.
 PT Prodn. of dicistronic DNA library used to make antibodies, etc. -
 PT includes forming 1st and 2nd PCR admixtures, subjecting them to
 PT PCR thermo-cycles, sepg. double stranded DNA, hybridising, etc.

PS Disclosure, Page 100, 143pp, English.
 CC This peptide linker sequence is used to increase the distance of an
 CC expressed IgG polypeptide from the surface membrane of E. coli, which
 CC results in decreased steric hindrance and competition of the preselected
 CC polypeptide with the lipopolysaccharide coat of E. coli. The IgG is
 CC expressed by a vector produced by a novel form of fusion PCR which
 CC enables fusion of heavy and light chains prior to vector ligation,
 CC avoiding the cumbersome separate cloning of fragments. This linker
 CC sequence moves the SpeI site, retains original IgG1 upper hinge region,
 CC retains original lamB sequence.
 SQ Sequence 34 AA;
 SQ 2 A; 0 R; 0 N; 2 D; 0 B; 1 C; 0 Q; 2 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 2 L; 5 K; 0 M; 1 F; 6 P; 5 S; 5 T; 0 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:26-PDT using FindSeq

Initial Score = 6 Optimized Score = 7 Significance = 3.37
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

```

      X      10      20 X
      TEFLSNYLTNVDDITLVPETLGR
      || | | || |
PKSCDKTHTPEKSTDKTHTSPAPAPPELLKSSFY
      10      20      30
  
```

13. US-08-249-182-11 (1-23)

R40101 Hib OMP P1-P2 hybrid peptide CP2-1P13.

ID R40101 standard; peptide; 54 AA.
 AC R40101;
 DT 04-FEB-1994 (first entry)
 DE Hib OMP P1-P2 hybrid peptide CP2-1P13.
 KW Haemophilus influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 1..16
 FT /label= C-P1
 FT Region 17..53
 FT /label= CHIBP2
 PN W09315205-A.
 PD 05-AUG-1993.
 PF 03-FEB-1993; CA0041.
 PR 03-FEB-1992; GB-002219.
 PA (CONN-) CONNAUGHT LAB LTD.
 PI Chong P, Kandil A, Klein MH, Sia C;
 DR WPI; 93-258681/32.
 PT Synthetic Haemophilus influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 PS Table 11; Page 59; 99pp; English.
 CC The sequences given in R40053-101 are peptide fragments derived from
 CC the Haemophilus influenzae type b (Hib) outer membrane proteins P1,
 CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise a immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.
 SQ Sequence 54 AA;
 SQ 4 A; 3 R; 5 N; 1 D; 0 B; 0 C; 1 Q; 3 E; 0 Z; 4 G; 1 H;
 SQ 1 I; 5 L; 6 K; 0 M; 2 F; 0 P; 2 S; 11 T; 0 W; 3 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:23-PDT using FindSeq

Initial Score = 6 Optimized Score = 7 Significance = 3.37
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

```

      X      10      20  X
      TEFLSNYLTVDDITLVPETLGR
      ||      |  |  ||  |
ARTRTTETGKGKVKTEKFEVKTIQDKRTLTLNNTANYTSQAHANLYGLNLNYSF
      10  X      20      30      X  40      50

```

14. US-08-249-182-11 (1-23)

R43686 Human kappa constant domain as encoded by pHCMV-KR

ID R43686 standard; Protein; 106 AA.
 AC R43686;
 DT 25-MAY-1994 (first entry)
 DE Human kappa constant domain as encoded by pHCMV-KR.
 KW Human; immunoglobulin; constant; region; humanised; P-selectin; light;
 KW blocking; antibody; heavy; chain; variable; murine; thrombotic disease;
 KW monoclonal; PB1.3; CDR; complementarity determining region; leukocyte;
 KW expression vector; coexpression; pHCMV-1748RHA-gammaCi-dhfr; epitope;
 KW pHCMV-1748RLA-KR-neo; PB1.3/Humanised version A; vascular endothelium;
 KW pHCMV-1747CH-gammaCi-neo; pHCMV-1747-CL-KR-neo; PB1.3 chimera;
 KW acute lung injury; ischaemia reperfusion injury; inflammation.
 OS Homo sapiens.
 PN WD9321956-A.
 PD 11-NOV-1993.
 PF 04-MAY-1993; U04274.
 PR 05-MAY-1992; US-880196.
 PA (CYTE-) CYTEL CORP.
 PI Chestnut RW, Paulson JC, Polley MJ;
 DR WPI; 93-368423/46.
 DR N-PSDB; 851548.
 PT Anti-P-selectin antibody for ischaemia acute lung injury treatment -
 PT useful to treat inflammation and pathological conditions of
 PT intercellular adhesion by competitive inhibition assays
 PS Example 10; Fig 10; 82pp; English.
 CC The sequences given in R43685-86 represent human immunoglobulin
 CC constant regions which were used in the production of the humanised
 CC P-selectin blocking antibody, along with the heavy and light chain
 CC variable region coding sequences of the murine monoclonal antibody
 CC PB1.3, given in R43687-88. The CDRs from PB1.3 heavy and light
 CC chains were substituted for the CDRs of human heavy and light chains.
 CC The humanised variable regions were inserted into expression vectors.
 CC By coexpression of appropriate combinations of heavy and light
 CC chains, several humanised antibodies can be expressed. Coexpression
 CC of pHCMV-1748RHA-gammaCi-dhfr and pHCMV-1748RLA-KR-neo gives rise
 CC to the PB1.3/Humanised version A. Coexpression of pHCMV-1747CH-
 CC gammaCi-neo and pHCMV-1747-CL-KR-neo gives rise to the PB1.3 chimera.
 CC These humanised antibodies selectively bind epitopes on P-selectin and
 CC block adhesion of leukocytes to the vascular endothelium. They may be
 CC used to treat inflammatory and thrombotic diseases and other
 CC pathological conditions involving P-selectin and antibodies to it, esp.
 CC acute lung injury and ischaemia reperfusion injury.
 SQ Sequence 106 AA;
 SQ 7 A; 2 R; 5 N; 5 D; 0 B; 3 C; 6 Q; 7 E; 0 Z; 4 G; 2 H;
 SQ 1 I; 8 L; 8 K; 0 M; 4 F; 5 P; 16 S; 8 T; 1 W; 4 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:42-PDT using FindSeq

Initial Score = 6 Optimized Score = 6 Significance = 3.37
 Residue Identity = 26% Matches = 6 Mismatches = 17
 Gaps = 0 Conservative Substitutions = 0

```

      X      10      20
      TEFLSNYLTVDDITLVPETLG

```


SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD
 10 20 30 40 50 X 60 70

X
 R

YEKHKVYACEVTHQGLSSPVTKSFNRGEC
 80 90 100

15. US-08-249-182-11 (1-23)

R41687 Undefined ORF1 encoded by plasmid pAH4611.

ID R41687 standard; Protein; 106 AA.
 AC R41687;
 DT 20-OCT-1993 (first entry)
 DE Undefined ORF1 encoded by plasmid pAH4611.
 KW Polymerase chain reaction; primer; PCR; amplify; murine; heavy; light;
 KW chain; variable; constant; region; anti-human; transferrin; receptor;
 KW antibody; brain; capillary; endothelial cell; conjugate; epilepsy;
 KW neuropharmaceutical; diagnostic; agent; tumour; AIDS; stroke;
 KW Parkinsons disease; Alzheimers disease.
 OS Synthetic.
 PN W09310819-A.
 PD 10-JUN-1993.
 PF 24-NOV-1992; U10206.
 PR 26-NOV-1991; US-800458.
 PA (ALKE-) ALKERMES INC.
 PI Friden PM;
 DR WPI; 93-196742/24.
 DR N-PSDB; Q43845.
 PT Antibody conjugates specific for transferrin receptor - used
 PT for diagnosis and treatment of cancer, AIDS and neurological
 PT disorders
 PS Disclosure; Fig 13H; 151pp; English.
 CC The sequences given in R41686-87 represent proteins encoded by the
 CC expression vector pAH4611. This vector was produced from the plasmid
 CC pAG4270. pAG4270 is the expression vector for the light chain
 CC variable region (VL) of the antibody 128.1 which was obtained by PCR
 CC with leader/J region priming (see also Q43842). The vector also
 CC contains an ampicillin resistance gene, a gpt (mycophenolic acid
 CC resistance) selected marker, an immunoglobulin H enhancer and an
 CC intron for V-constant region splicing. Transcription of the CH gene
 CC is from the VH promoter of the murine 27.44 gene. The cloning of
 CC the 128.1 VL region was accomplished in two stages with the production
 CC of plasmid pSV4271 as an intermediate vector which lacks the promoter
 CC region. This plasmid was used in conjunction with pAH4602 in the
 CC production of the chimeric antibody 128.1. 128.1 is an anti-human
 CC transferrin receptor antibody which binds to the transferrin receptor
 CC on brain capillary endothelial cells. This antibody may be used in a
 CC conjugate in which it is linked to a neuropharmaceutical or diagnostic
 CC agent. The conjugate may be used to treat or prevent neurological
 CC disorders eg. brain tumours, AIDS, stroke, epilepsy, Parkinsons and
 CC Alzheimers disease. It may also be used for diagnostic methods.
 SQ Sequence 106 AA;
 SQ 7 A; 2 R; 5 N; 5 D; 0 B; 3 C; 6 Q; 7 E; 0 Z; 4 G; 2 H;
 SQ 1 I; 8 L; 8 K; 0 M; 4 F; 5 P; 16 S; 8 T; 1 W; 4 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:14-PDT using FindSeq

Initial Score = 6 Optimized Score = 6 Significance = 3.37
 Residue Identity = 26% Matches = 6 Mismatches = 17
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNVLTNVDDITLVPETLG

DR WPI; 93-272183/34.
 PT New purified enterotoxin receptor protein - used to develop
 PT prods. for treating abnormal conditions caused by bacterially
 PT released enterotoxin, partic. diarrhoea
 PS Disclosure; Fig 3; 26pp; English.
 CC The sequences given in R38862-63 represent the guanylyl cyclases,
 CC GC-A and GC-B, which binds heat stable enterotoxin. These proteins
 CC are enterotoxin receptors which may be used as a therapeutic to control
 CC intestinal fluid permeation as well as abnormal conditions caused
 CC by bacterially released enterotoxin. The binding domain of the
 CC proteins, or antibodies to the proteins, can be used to eliminate
 CC diarrhoea. The proteins may be used to isolate ligands and to screen
 CC for antagonists of toxin binding. This sequence is given as it is
 CC represented in the specification.
 SQ Sequence 1025 AA;
 SQ 82 A; 69 R; 39 N; 54 D; 0 B; 15 C; 40 Q; 64 E; 0 Z; 78 G; 28 H;
 SQ 49 I; 122 L; 39 K; 21 M; 46 F; 53 P; 56 S; 53 T; 16 W; 37 Y; 64 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 GETVQAEAFDSVTIYFSDIVGFTALSAESTPMQVVTLNDLYTCFDAIIDNFVYK VETIGDAYMVVSGLPG
 830 840 850 860 870 880 X 890

 RNGQRHAFEIARMALALLDAVSSFRIRHRPHDQL
 900 910 920 930

9. US-08-249-182-2 (1-6)

R38862 GC-A.

ID R38862 standard; Protein; 1029 AA.
 AC R38862;
 DT 08-FEB-1994 (first entry)
 DE GC-A.
 KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
 KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
 KW binding domain; antibody; diarrhoea; ligand; antagonist.
 OS Rattus rattus.
 PN US5237051-A.
 PD 17-AUG-1993.
 PF 06-DEC-1990; 623033.
 PR 06-DEC-1990; US-623033.
 PA (UYVA-) UNIV VANDERBILT.
 PI Garbers DL, Schulz S;
 DR WPI; 93-272183/34.
 PT New purified enterotoxin receptor protein - used to develop
 PT prods. for treating abnormal conditions caused by bacterially
 PT released enterotoxin, partic. diarrhoea
 PS Disclosure; Fig 3; 26pp; English.
 CC The sequences given in R38862-63 represent the guanylyl cyclases,
 CC GC-A and GC-B, which binds heat stable enterotoxin. These proteins
 CC are enterotoxin receptors which may be used as a therapeutic to control
 CC intestinal fluid permeation as well as abnormal conditions caused
 CC by bacterially released enterotoxin. The binding domain of the
 CC proteins, or antibodies to the proteins, can be used to eliminate
 CC diarrhoea. The proteins may be used to isolate ligands and to screen
 CC for antagonists of toxin binding.
 SQ Sequence 1029 AA;
 SQ 74 A; 73 R; 37 N; 56 D; 0 B; 16 C; 40 Q; 72 E; 0 Z; 75 G; 23 H;

FT /note= "1 N-linked glycosylation site"
 FT Region 552..766
 FT /label= C-terminal region of extracellular
 FT domain
 FT /note= "1 N-linked glycosylation site & 1
 FT catalytic site"
 FT Active_site 627..631
 FT /label= active site of serine protease/esterase
 FT /note= "fits the consensus sequence GXSGXG"
 PN W09316102-A.
 PD 19-AUG-1993.
 PF 09-APR-1992; U02892.
 PR 06-FEB-1992; US-832211.
 PA (DAND) DANA FARBER CANCER INST INC.
 PI Morimoto C, Schlossman SF, Tanaka T;
 DR WPI; 93-272827/34.
 DR N-PSDB; 046089.
 PT Polypeptide fragments of CD26 - are capable of disrupting binding
 PT of CD45 and CD26 and thus interfering with T-cell activation
 PS Disclosure; pages 39-43; 73pp; English.
 CC C26 is a human T cell activation antigen originally identified by
 CC its reactivity with the MAbs Ta1. C26 cDNA library was constructed
 CC from human PHA-activated T cells using the CDM7vector. The hydrophobic
 CC N-terminal of the predicted CD26 polypeptide has the characteristics
 CC of a signal sequence of the type II membrane protein, which is
 CC reinforced by the observation that potential N-glycosylation sites
 CC are located in the carboxy side of the hydrophobic core. Therefore
 CC the N-terminal 6 AAs are predicted to be cytoplasmic, the next 22
 CC AAs are predicted to transverse the cytoplasmic membrane, and the
 CC 738 C-terminal AAs constitute the predicted extracellular domain.
 SQ Sequence 766 AA;
 SQ 40 A; 30 R; 40 N; 46 D; 0 B; 12 C; 30 Q; 40 E; 0 Z; 43 G; 19 H;
 SQ 49 I; 62 L; 40 K; 15 M; 31 F; 29 P; 64 S; 50 T; 21 W; 56 Y; 49 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:24-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 VLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQMLPPHFDKSKKYPLLLDVYAGPCSQKADTVFRLNWA
 500 510 520 530 540 X 550 560
 TYLASTENIIVASFDRGSGYQGDKIMHAINRRL
 570 580 590

8. US-08-249-182-2 (1-6)
 R38863 GC-B.

ID R38863 standard; Protein; 1025 AA.
 AC R38863;
 DT 08-FEB-1994 (first entry)
 DE GC-B.
 KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
 KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
 KW binding domain; antibody; diarrhoea; ligand; antagonist.
 OS Rattus rattus.
 PN US5237051-A.
 PD 17-AUG-1993.
 PF 06-DEC-1990; 623033.
 PR 06-DEC-1990; US-623033.
 PA (UYVA-) UNIV VANDERBILT.

CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                ||||
GETVQAEAFDSVTIYFSDIVGFTALSAESTPMQVVTLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGLPV
      840      850      860      870      880 X      890      900

RNGQLHAREVARMALALLDAVRSFRIRHRPQEQ
      910      920      930
```

10. US-08-249-182-2 (1-6)

R24102 Marek's disease virus MD20 polypeptide.

ID R24102 standard; Protein; 1074 AA.
AC R24102;
DT 14-NOV-1992 (first entry)
DE Marek's disease virus MD20 polypeptide.
KW Antibodies; vaccine; recombinant; poultry; passive immunotherapy;
KW diagnostic immunoassay; anti-idiotypic; antigen.
OS Marek's disease virus.
PN EP-486106-A.
PD 20-MAY-1992.
PF 13-NOV-1991; 202947.
PR 16-NOV-1990; US-615211.
PA (ALKU) AKZO NV.
PI Morgan RW;
DR WPI; 92-168713/21.
DR N-PSDB; 024789.
PT DNA encoding Marek's disease virus polypeptides MD18 and MD20 -
PT and antibodies and vaccine useful for the protection of poultry
PT against MDV infection
PS Claim 8; Page 18; 31pp; English.
CC The protein sequence of MDV MD20 was deduced from the DNA sequence
CC obtd. by screening a lambda EMBL 3 library made by infecting chicken
CC embryo fibroblasts with a tissue-culture adapted passage of Marek's
CC disease virus (MDV) strain GA, and incubating until a 90 percent
CC cytopathic effect had developed. Vectors and host cells contg. the
CC MDV MD20 gene and MDV polypeptides can be used in a vaccine to protect
CC poultry against Marek's disease. Antibodies or antiserum raised by
CC the polypeptides may be used in passive immunotherapy, diagnostic
CC immunoassays and in the generation of anti-idiotypic antibodies for
CC use in a test kit for Marek's disease. The vaccine may also contain
CC immunogens related to other poultry pathogens, e.g. infectious
CC bronchitis-virus, Newcastle disease-virus or infectious bursal
CC disease-virus to produce a multivalent vaccine.
CC See also R24102.
SQ Sequence 1074 AA;
SQ 72 A; 67 R; 49 N; 66 D; 0 B; 30 C; 25 Q; 56 E; 0 Z; 68 G; 27 H;
SQ 80 I; 109L; 49 K; 24 M; 49 F; 43 P; 92 S; 52 T; 8 W; 41 Y; 67 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:06-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                ||||
SGMILECSNTHSRGGPMIVKRLALVSADSRAGGIGPANMLMGIDSAIDGPLPVYRVGMSKGRQAFVLMTE
```

11. US-08-249-182-2 (1-6)

R38861 GC-C.

ID R38861 standard; Protein; 1075 AA.
AC R38861;
DT 08-FEB-1994 (first entry)
DE GC-C.
KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
KW binding domain; antibody; diarrhoea; ligand; antagonist.
OS Rattus rattus.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /note= "Signal peptide"
FT Protein 23..1075
FT /note= "Mature GC-C"
FT Modified_site 31..33
FT /note= "N-linked glycosylation site"
FT Modified_site 74..76
FT /note= "N-linked glycosylation site"
FT Modified_site 78..80
FT /note= "N-linked glycosylation site"
FT Modified_site 187..189
FT /note= "N-linked glycosylation site"
FT Modified_site 194..196
FT /note= "N-linked glycosylation site"
FT Modified_site 306..308
FT /note= "N-linked glycosylation site"
FT Modified_site 356..358
FT /note= "N-linked glycosylation site"
FT Modified_site 401..403
FT /note= "N-linked glycosylation site"
FT Domain 433..453
FT /note= "Transmembrane domain"
PN US5237051-A.
PD 17-AUG-1993.
PF 06-DEC-1990; 623033.
PR 06-DEC-1990; US-623033.
PA (UYVA-) UNIV VANDERBILT.
PI Garbers DL, Schulz S;
DR WPI; 93-272183/34.
DR P-PSDB; R38861.
PT New purified enterotoxin receptor protein - used to develop
PT prods. for treating abnormal conditions caused by bacterially
PT released enterotoxin, partic. diarrhoea
PS Claim 2; Fig 1; 26pp; English.
CC This sequence represents guanylyl cyclase, GC-C, which binds heat
CC stable enterotoxin. The DNA encoding this protein was isolated from
CC rat small intestinal mucosa polyA+ RNA by PCR. This protein is an
CC enterotoxin receptor which may be used as a therapeutic to control
CC intestinal fluid permeation as well as abnormal conditions caused
CC by bacterially released enterotoxin. The binding domain of the
CC protein, or antibodies to the protein, can be used to eliminate
CC diarrhoea. The protein may be used to isolate ligands and to screen
CC for antagonists of toxin binding.
SQ Sequence 1075 AA;
SQ 42 A; 64 R; 48 N; 68 D; 0 B; 20 C; 32 Q; 73 E; 0 Z; 55 G; 23 H;
SQ 57 I; 118L; 68 K; 32 M; 52 F; 43 P; 75 S; 71 T; 11 W; 48 Y; 75 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

ID R40916 standard; Protein; 593 AA.
AC R40916;
DT 05-FEB-1994 (first entry)
DE Sequence of a CD26 fragment lacking a portion of the carboxy
DE terminal region
KW Human T cell activation antigen; monoclonal antibody Tal; CD26.
OS Synthetic.
PN W09316102-A.
PD 19-AUG-1993.
PF 09-APR-1992; U02892.
PR 06-FEB-1992; US-832211.
PA (DAND) DANA FARBER CANCER INST INC.
PI Morimoto C, Schlossman SF, Tanaka T;
DR WPI; 93-272827/34.
PT Polypeptide fragments of CD26 - are capable of disrupting binding
PT of CD45 and CD26 and thus interfering with T-cell activation
PS Example; Pages 46-48; 73pp; English.
CC C26 is a human T cell activation antigen originally identified by
CC its reactivity with the MAb Tal. C26 cDNA library was constructed
CC from human PHA-activated T cells using the CDM7 vector.
CC Fragments of CD26 can be prep'd in the following manner.
CC CD26 XbaI-SphI cDNA fragment is ligated to the vector
CC R_cSR-alpha-26 XbaI-HindIII DNA fragment and the linker Q46092.
CC The linker introduces an in-frame stop codon that results in the
CC deletion of the segment of CD26 from AA 594 to the carboxy
CC terminus of the wild-type protein. This deletion mutant, shown
CC in R40916, lacks the putative catalytic site of CD26 and has a new
CC carboxy terminus given in R40917.
SQ Sequence 593 AA;
SQ 28 A; 22 R; 34 N; 35 D; 0 B; 10 C; 21 Q; 31 E; 0 Z; 30 G; 11 H;
SQ 39 I; 54 L; 34 K; 9 M; 22 F; 25 P; 49 S; 41 T; 17 W; 46 Y; 35 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:24-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

                X  X
                PLDVYK
                IIII
VLEDNSALDKMLGNVQMPSKKLDFIILNETKFWYQMI LPPHFDKSKKYPLLLDVYAGPCSQKADTVFRLNWA
   500       510       520       530       540 X   550       560

TYLASTENIIVASFDRGSGYQGDKIMHA
   570       580       590

```

7. US-08-249-182-2 (1-6)

R40909 Sequence encoded by human CD26 cDNA.

ID R40909 standard; Protein; 766 AA.
AC R40909;
DT 05-FEB-1994 (first entry)
DE Sequence encoded by human CD26 cDNA.
KW Human T cell activation antigen; monoclonal antibody Tal.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 7..28
FT /label= hydrophobic
FT Region 29..323
FT /label= N-terminal glycosylated region of
FT extracellular domain
FT /note= "8 sites for N-linked glycans"
FT Region 324..551
FT /label= Cysteine rich region of extracellular

Gaps = 0 Conservative Substitutions = 0

```

          X   X
          PLDVYK
          |||
DSGEYKCGHQGVNESEPVYLEVFSWLLLGASAEVVMEGGPLFLRCHGWRNWDVYKVYYKDGEALKYWYEN
 90      100      110      120      130      140 X      150

HNISITNATVEDSGTYVCTGKVMQLDYESEPLNI
160      170      180      190

```

5. US-08-249-182-2 (1-6)

R42635 Human interferon receptor.

ID R42635 standard; Protein; 557 AA.
AC R42635;
DT 20-APR-1994 (first entry)
DE Human interferon receptor.
KW IFN-R; extracellular domain; monoclonal antibody; viral infection;
KW cell proliferation; allograft rejection; systemic lupus erythematosus;
KW psoriasis; multiple sclerosis; Behcet's Disease; aplastic anaemia;
KW immunodeficiency; measles virus; interferon-alpha-beta.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..436
FT /label= extracellular_domain
FT /note= "soluble, immunogenic form of IFN-R"
PN EP-563487-A.
PD 06-OCT-1993.
PF 31-MAR-1992; 400902.
PR 31-MAR-1992; EP-400902.
PA (EUBI-) LAB EURO BIOTECHNOLOGIE SA.
PI Benoit P, Maguire D, Meyer F, Plavec I, Tovey MG;
DR WPI; 93-312951/40.
DR P-PSDB; R42635.
PT Monoclonal antibody to human interferon type-I receptor - having
PT neutralising activity against human type I interferon, used for
PT therapy and diagnosis
PS Disclosure; Fig 3; 21pp; English.
CC Monoclonal antibodies produced against soluble forms of the human
CC interferon alpha-beta receptor based on the full-length human IFN-R
CC sequence are claimed. The antibodies are useful for treatment and
CC prophylaxis of disorders involving cell proliferation and/or viral
CC infection.
SQ Sequence 557 AA;
SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
SQ 43 I; 44 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 42 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:33-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

          X   X
          PLDVYK
          |||
IYIGAPKQSGNTPVVIQDYPLIYEIIFWENTSNAERKIIIEKKTDTVLPNLKPLTVYCVKARAHTMDEKLNKSS
350      360      370      380      390      400 X      410

VFSDAVCEKTKPGNTSKIWLIVGICIALFALPFV
420      430      440      450

```

6. US-08-249-182-2 (1-6)

R40916 Sequence of a CD26 fragment lacking a portion of t

PS Labelled antibody to N terminus
 CC Antibodies were raised against two synthetic peptides that
 CC correspond to the N-terminal (R36681) and C-terminal (R36680)
 CC of guinea pig VPF (designated N-IgG and C-IgG, respectively).
 SQ Sequence 25 AA;
 SQ 2 A; 2 R; 0 N; 1 D; 0 B; 1 C; 1 Q; 3 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 0 L; 3 K; 2 M; 1 F; 2 P; 1 S; 0 T; 0 W; 2 Y; 3 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:00-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 APMAEGEQKPREVVKFMVYKRSYC
 10 X 20

4. US-08-249-182-2 (1-6)

R26064 Human FcERI alpha-subunit and IL-2 hybrid protein.

ID R26064 standard; Protein; 235 AA.
 AC R26064;
 DT 02-FEB-1993 (first entry)
 DE Human FcERI alpha-subunit and IL-2 hybrid protein.
 KW High affinity Fc immunoglobulin E receptor; IgE;
 KW antibody; interleukin-2; histamine release; allergy.
 DS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 26..201
 FT /label= human_FcERI_alpha-subunit
 FT /note= "epitope recognised by new MAbs"
 PN EP-499112-A.
 PD 19-AUG-1992.
 PF 03-FEB-1992; 101732.
 PR 11-FEB-1991; US-653936.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Chizzonite RA, Hakimi J, Kochan JP;
 DR WPI; 92-277871/34.
 PT Monoclonal antibodies bind to alpha sub-unit of Fc IgE receptor -
 PT for treatment and prevention of IgE induced allergic diseases,
 PT also for measuring alpha sub-unit and IgE levels in biological
 PT fluids
 PS Disclosure; Page 8; 30pp; English.
 CC This is a preferred protein for use in generating the
 CC monoclonal antibodies of the invention. The protein comprises
 CC an epitope of the human FcERI alpha-subunit to which the cytoplasmic
 CC and transmembrane regions of the IL-2 receptor have been fused.
 CC (Cytoplasmic and transmembrane regions from receptors other IL-
 CC 2 receptor can be also used). The specification includes a
 CC nucleotide coding sequence which is a preferred fusion gene (see
 CC Q27267); the polypeptide which is decoded from that fusion gene
 CC differs from the amino acid sequence R26064 as follows: amino acids
 CC 5-7 are Arg-Ile-Leu (not Met-Glu-Ser), amino acid 209 is Cys (not
 CC Lys), amino acid 229 is Ser (not Arg), Arg233 is absent and an
 CC additional C-terminal amino acid (Phe) is present.
 SQ Sequence 235 AA;
 SQ 13 A; 9 R; 16 N; 7 D; 0 B; 5 C; 10 Q; 18 E; 0 Z; 12 G; 4 H;
 SQ 9 I; 24 L; 14 K; 3 M; 10 F; 11 P; 18 S; 12 T; 8 W; 11 Y; 21 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:22-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2

PLDVYK
X X

2. US-08-249-182-2 (1-6)

R13381 Vascular permeability factor (1-21).

ID R13381 standard; peptide; 21 AA.
AC R13381;
DT 22-OCT-1991 (first entry)
DE Vascular permeability factor (1-21).
KW VPF; antibody.
OS Synthetic.
PN US5036003-A.
PD 30-JUL-1991.
PF 02-SEP-1988; 240780.
PR 21-AUG-1987; US-087739.
PR 02-SEP-1988; US-240780.
PA (MONS) MONSANTO CO.
PI Olander JV, Connolly DT, Adams SP, Feder J.
DR WPI; 91-245462/33.
PT Antibodies to vascular permeability factor peptides - which block
PT permeability-enhancing activity and growth-promoting activity of
PT vascular permeability factor.
PS Claim 2; Page 5; 6pp; English.
CC The amino acid sequence encodes a synthetic peptide fragment of VPF
CC prepared by solid phase synthesis which is used to stimulate an
CC antibody response. This response blocks the permeability-enhancing
CC and growth promoting activity of VPF used by tumours to increase
CC their nutrient supply. The Abs are also useful as immunoabsorbents
CC for VPF isolation and in assays for VPF. See also R13380.
SQ Sequence 21 AA;
SQ 2 A; 1 R; 0 N; 1 D; 0 B; 0 C; 1 Q; 3 E; 0 Z; 1 G; 0 H;
SQ 0 I; 0 L; 3 K; 2 M; 1 F; 2 P; 0 S; 0 T; 0 W; 1 Y; 3 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:33-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.88
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||||
APMAEGEQKPREVVKFMDVYK
10 X 20

3. US-08-249-182-2 (1-6)

R36681 Guinea pig VPF N-terminal.

ID R36681 standard; peptide; 25 AA.
AC R36681;
DT 23-AUG-1993 (first entry)
DE Guinea pig VPF N-terminal.
KW Vascular permeability factor; effusion; malignancy; antibody.
OS Cavia porcellus.
PN W09308473-A.
PD 29-APR-1993.
PF 21-OCT-1992; U09068.
PR 24-OCT-1991; US-782350.
PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
PI Dvorak HF, Yeo K, Yeo T;
DR WPI; 93-152625/18.
PT Immunoassay for detecting vascular permeability (VPF) factor in
PT effusions - comprises detecting VPF in sample using immobilised
PT antibody to C-terminus of VPF having specified sequence and 2nd
PT

22. P98438	Sequence of C. trachomatis se	14	3	3	2.91	0
23. P98432	Sequence of C. trachomatis se	14	3	3	2.91	0
24. P98464	Sequence of C. trachomatis se	14	3	3	2.91	0
25. P98428	Sequence of C. trachomatis se	14	3	3	2.91	0
26. P98456	Sequence of C. trachomatis se	14	3	3	2.91	0
27. P98420	Sequence of C. trachomatis se	14	3	3	2.91	0
28. P98416	Sequence of C. trachomatis se	14	3	3	2.91	0
29. P98412	Sequence of C. trachomatis se	14	3	3	2.91	0
30. P98452	Sequence of C. trachomatis se	14	3	3	2.91	0
31. P98448	Sequence of C. trachomatis se	14	3	3	2.91	0
32. P98468	Sequence of C. trachomatis se	14	3	3	2.91	0
33. P98460	Sequence of C. trachomatis se	14	3	3	2.91	0
34. P93297	Chlamydia trachomatis serovar	14	3	3	2.91	0
35. P93291	Sequence of Chlamydia trachom	14	3	3	2.91	0
36. P91447	Sequence of Chlamydia trachom	14	3	3	2.91	0
37. R43907	HIV-1 RF gp120 monoclonal ant	15	3	3	2.91	0
38. R34502	Immunogenic AHAS peptide #4.	15	3	3	2.91	0
39. R44830	Human fibrin beta-chain inter	16	3	3	2.91	0
40. P93015	38-57 region of equine follic	20	3	3	2.91	0

1. US-08-249-182-2 (1-6)

R37444 Autotaxin peptide ATX 19.

ID R37444 standard; peptide; 7 AA.
AC R37444;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 19.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
FH Key Location/Qualifiers
FT Modified_site 2
FT /note= "potentially glycosylated residue"
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 19. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 7 AA;
SQ 0 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 0 T; 0 W; 1 Y; 1 V;
SQ 1 Others;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.85
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||||

SCORE	0	1	1	2	2	3	3	4	4	5
STDEV	1		2		3		4			

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.03

Times:	CPU	Total Elapsed
	00:00:09.94	00:00:16.00

Number of residues:	482836
Number of sequences searched:	5543
Number of scores above cutoff:	2034

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
1. R37444	Autotaxin peptide ATX 19.	7	5	5	4.85	0
**** 3 standard deviations above mean ****						
2. R13381	Vascular permeability factor	21	4	4	3.88	0
3. R36681	Guinea pig VPF N-terminal.	25	4	4	3.88	0
4. R26064	Human FcERI alpha-subunit and	235	4	4	3.88	0
5. R42635	Human interferon receptor.	557	4	4	3.88	0
6. R40916	Sequence of a CD26 fragment 1	593	4	4	3.88	0
7. R40909	Sequence encoded by human CD2	766	4	4	3.88	0
8. R38863	GC-B.	1025	4	4	3.88	0
9. R38862	GC-A.	1029	4	4	3.88	0
10. R24102	Marek's disease virus MD20 po	1074	4	4	3.88	0
11. R38861	GC-C.	1075	4	4	3.88	0
12. P60243	Sequence encoding the entire	2179	4	4	3.88	0
**** 2 standard deviations above mean ****						
13. P30125	Sequence of immunoregulatory	6	3	3	2.91	0
14. P71314	Sequence of fibrin immunogen	7	3	3	2.91	0
15. R44837	Human fibrin beta-chain N-ter	8	3	3	2.91	0
16. P82686	Human fibrin beta chain N-ter	8	3	3	2.91	0
17. R44838	Human fibrin beta-chain pepti	12	3	3	2.91	0
18. R44829	Human fibrin beta-chain N-ter	12	3	3	2.91	0
19. R28629	N-terminal human fibrin pepti	12	3	3	2.91	0
20. R42343	C-erbB-2 C-terminal peptide.	13	3	3	2.91	0
21. P90379	Sequence of antigenic epitope	14	3	3	2.91	0

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                ||||
KGIVEPELYEEVTIYFSDIVGFTTICKYSTPMEVVDMNDIYKSFDDGIVDHHDVYKVETIGDAYVVASGLPM
  820      830      840      850      860  X  870      880

RNGNRHVAVDISKMALDILSFMGTFELEHLPLPLPV
  890      900      910

```

12. US-08-249-182-2 (1-6)

P60243 Sequence encoding the entire genomic RNA of human

ID P60243 standard; Protein; 2179 AA.
 AC P60243;
 DT 07-AUG-1991 (first entry)
 DE Sequence encoding the entire genomic RNA of human rhinovirus
 KW Monoclonal antibody; MAB; HRV; vaccine; ss.
 OS Human rhinovirus.
 FH Key Location/Qualifiers
 FT Region 1..69
 FT /label= VP4 structural protein
 FT Region 70..331
 FT /label= VP2 structural protein
 FT Region 332..567
 FT /label= VP3 structural protein
 FT Region 568..856
 FT /label= VP1 structural protein
 FT Region 857..1002
 FT /label= 3B protein
 FT Region 1003..1099
 FT /label= 5B protein
 FT Region 1100..1429
 FT /label= X protein
 FT Region 1430..1514
 FT /label= protein 1B
 FT Region 1515..1537
 FT /label= protein VPg
 FT Region 1538..1719
 FT /label= Protease
 FT Region 1720..2179
 FT /label= Replicase
 PN EP-169146-A.
 PD 22-JAN-1986.
 PF 17-JUL-1985; 401465.
 PR 20-JUL-1984; US-632785.
 PR 10-APR-1985; US-721735.
 PA (MERI) MERCK & CO INC.
 PI Colonna RJ, Mitzutani S;
 DR WPI; 86-022809/04.
 DR N-PSDB; N60194.
 PT New DNA encoding the entire genomic RNA of human rhinovirus 14 -
 PT and monoclonal antibodies which block attachment or neutralise
 PT infectivity of rhinovirus.
 PS Example 11; Page 22-39; 80pp; English.
 CC Sequence may be used for the manufacture of hybridoma cells
 CC expressing the HRV or fragments thereof. The fusion products may be
 CC used in immunisation, or to raise MAbs for passive treatment of HRV
 CC infection.
 SQ Sequence 2179 AA;
 SQ 110A; 82 R; 114N; 117D; 0 B; 43 C; 81 Q; 103E; 0 Z; 146G; 53 H;
 SQ 144I; 196L; 137K; 53 M; 85 F; 121P; 158S; 172T; 24 W; 85 Y; 155V;

Initial Score = 4 Optimized Score = 4 Significance = 3.88
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK

II II
 FAGSSLIQSICNTHHIFRDEIYVVEGGMPSGCGTSIFNSHINNIIIRTLILDAYKCIDLDKILAYGDDL
 1980 1990 2000 2010 2020 2030 X 2040

IVSYPYELDPQVLATLGKNYGLTITPPDKSETFT
 2050 2060 2070 2080

13. US-08-249-182-2 (1-6)

P30125 Sequence of immunoregulatory peptides I, II, XIII,

ID P30125 standard; peptide; 6 AA.

AC P30125;

DT 03-AUG-1992 (first entry)

DE Sequence of immunoregulatory peptides I, II, XIII, XIV, and XV.

KW Pharmaceutical ; immunoregulation; peptide; lymphocyte activity;

KW antibody.

FH Key Location/Qualifiers

FT Misc_difference 1

FT /label= pyro-E

FT Misc_difference 2..4

FT /label= I

FT Misc_difference 2..5

FT /label= II

FT Misc_difference 1..4

FT /label= XIII

FT Misc_difference 1..5

FT /label= XIV

FT Misc_difference 1..6

FT /label= XV

PN EP--67425-A.

PD 22-DEC-1982.

PF 11-JUN-1982; 387655.

PR 12-JUN-1981; HU-001755.

PA (RICT) RICHTER GEDEON VEGY.

PI Kisfaludy L, Nyeki D, Schoen I, Denes L, Ember J, Hajos G,

PI Szporny L, Szende B;

DR WPI; 83-00264K/01.

PT Protected tri; tetra; and penta;peptide(s) - for influencing

PT immuno-regulation e.g. by affecting lymphocyte activity, antibody

PT prodn., defence mechanisms

PS Claim 1; Page 45; 51pp; German.

CC The inventors claim 15 (peptides I-XV) opt. protected peptides and
 CC their salts, complexes, amides and 1-5C alkyl esters (see P30125-35).

CC The peptides (I)-(XV) are pharmaceuticals, esp. for influencing
 CC immunoregulation, with an effect on e.g. lymphocyte activity,
 CC antibody prodn. and cells which produce specific defences.

SQ Sequence 6 AA;

SQ 0 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;

SQ 0 I; 0 L; 1 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

CC Retrieved by shears on Wed 21 Sep 94 11:58:02-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 2.91
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK

PT harley JB;
 DR WPI; 93-351658/44.
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 PS Claim 1; Page 30; 43pp; English.
 CC The sequences given in R43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 SQ Sequence 10 AA;
 SQ 3 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 2 K; 0 M; 1 F; 0 P; 1 S; 1 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:35-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 STKQAAFKAV
 X 10

7. US-08-249-182-3 (1-5)

R07105 Hepatitis B Virus PreS1 epitope (I).

ID R07105 standard; peptide; 11 AA.
 AC R07105;
 DT 24-JAN-1991 (first entry)
 DE Hepatitis B Virus PreS1 epitope (I).
 KW Monoclonal antibodies; Hepatitis B Virus PreS1; PreS2; epitope;
 KW immunoassay; vaccine.
 OS Synthetic.
 PN EP-389983-A.
 PD 03-OCT-1990.
 PF 23-MAR-1990; 105550.
 PR 31-MAR-1989; US-332014.
 PR 25-MAY-1989; US-357708.
 PA (ABBO) ABBOTT LABORATORIES.
 PI Minns LT, Floreani MF, Eble KS, Rosenlof RV, Tyner JD;
 PI Witters E;
 DR WPI; 90-298968/40.
 PT Monoclonal antibodies to hepatitis B virus preS1 and preS2
 PT epitope(s) - use in immunoassays antibody assays, sub-typing HBV,
 PT vaccines, etc.
 PS Claim 1; Page 20; 22pp; English.

CC immunogens, followed by fusion. The MAb were then gp'd. by binding and
 CC inhibition studies using peptides corresp. to epitopes specific for
 CC PreS1, as represented here.
 CC The antibody binds to the PreS1 protein and not to the M protein.
 CC The MAbs can be used to develop specific and sensitive immunoassays
 CC to detect HBV and HB surface antigen in samples. They can also be
 CC used in immunoassays and to detect, characterise and isolate
 CC epitopal sites in PreS1 that may be useful components of
 CC subunit HBV vaccines. They have a potential as a prophylactic and
 CC curative agent. The Ab may further be used to map the hepatocyte
 CC binding region of the L protein and for intracellular and cell
 CC surface staining of HBV infected hepatocytes and experimentally
 CC transfected hepatoma cell lines.
 CC See also R07106-07.
 SQ Sequence 11 AA;
 SQ 2 A; 0 R; 1 N; 1 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 0 I; 1 L; 0 K; 0 M; 1 F; 1 P; 1 S; 0 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:09-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 HQLDPAFGANS
 X 10

8. US-08-249-182-3 (1-5)

R43438 Ro/SSA epitope 106.

ID R43438 standard; peptide; 12 AA.
 AC R43438;
 DT 12-MAY-1994 (first entry)
 DE Ro/SSA epitope 106.
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 PN W09321223-A.
 PD 28-OCT-1993.
 PF 13-APR-1993; U03484.
 PR 13-APR-1992; US-867819.
 PA (OKLA) UNIV OKLAHOMA STATE.
 PI Harley JB;
 DR WPI; 93-351658/44.
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 PS Claim 1; Page 30; 43pp; English.
 CC The sequences given in R43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin

phosphoprotein which associates with RNA polymerase III transcripts.
 Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 U5 RNA. Anti-Sm antibodies may be directed against one or a
 combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 used for preventing, treating or screening autoimmune disorders,
 especially SLE or Sjogrens syndrome (SS). They bind to a human
 autoantibody and may therefore be used as vaccines.
 Sequence 12 AA;
 3 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 0 H;
 1 I; 0 L; 2 K; 0 M; 1 F; 0 P; 2 S; 1 T; 0 W; 0 Y; 1 V;
 Retrieved by shears on Wed 21 Sep 94 11:59:35-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 ISTKQAAFKAVS
 X 10

9. US-08-249-182-3 (1-5)
 R43392 La/SSb epitope 139.

ID R43392 standard; peptide; 13 AA.
 AC R43392;
 DT 12-MAY-1994 (first entry)
 DE La/SSb epitope 139.
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 PN WD9321223-A.
 PD 28-OCT-1993.
 PF 13-APR-1993; U03484.
 PR 13-APR-1992; US-867819.
 PA (OKLA) UNIV OKLAHOMA STATE.
 PI Harley JB;
 DR WPI; 93-351658/44.
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 PS Claim 1; Page 30; 43pp; English.
 CC The sequences given in R43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be

US-08-249-182-3 (1-5)
CC especially SLE or Sjogrens syndrome (SS). They bind to a human
CC autoantibody and may therefore be used as vaccines.
SQ Sequence 13 AA;
SQ 1 A; 2 R; 1 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 1 H;
SQ 1 I; 1 L; 2 K; 1 M; 1 F; 0 P; 0 S; 1 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:34-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 60% Matches = 3 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
NIQMRRTLHKAFK
10 X

10. US-08-249-182-3 (1-5)

P91772 Mas oncongene angiotensin receptor peptide.

ID P91772 standard; protein; 16 AA.
AC P91772;
DT 14-MAY-1990 (first entry)
DE Mas oncongene angiotensin receptor peptide.
KW Antibody; angiotensin; mas; hypertension;
PN W08911657-A.
PD 30-NOV-1989.
PF 25-MAY-1989; G00583.
PR 25-MAY-1988; GB-012354.
PR 25-MAY-1988; GB-012353.
PA (CAMB) Cambridge Res Bioch.
PI Sheppard PW, Varndell IM;
DR WPI; 89-370824/50.
PT Detection of (pre-)hypertensive and/or (pre-)cancerous states -
PT and new peptide(s) used in raising antibodies at angiotensin
PT receptor useful in immunoassay, and treatment with angiotensin
PT antagonist.
PS Claim 3; Page 22; 32pp; English.
CC Antibodies raised against the peptide can be used for detection and
CC treatment of hypertension, tumours and other neoplasms, and prevention of
CC ectopic hormone prodn.
CC See also P91707 and P91771.
SQ Sequence 16 AA;
SQ 1 A; 3 R; 1 N; 2 D; 0 B; 1 C; 2 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 2 K; 1 M; 1 F; 1 P; 0 S; 0 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:56:50-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 75% Matches = 3 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
RAFKDEMQRPRQKDNC
X X 10

11. US-08-249-182-3 (1-5)

R43391 La/SSb epitope 136.

ID R43391 standard; peptide; 18 AA.
AC R43391;
DT 12-MAY-1994 (first entry)

DE La/SSB epitope 158.
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 PN W09321223-A.
 PD 28-OCT-1993.
 PF 13-APR-1993; U03484.
 PR 13-APR-1992; US-867819.
 PA (OKLA) UNIV OKLAHOMA STATE.
 PI Harley JB;
 DR WPI; 93-351658/44.
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 PS Claim 1; Page 30; 43pp; English.
 CC The sequences given in R43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 SQ Sequence 18 AA;
 SQ 1 A; 2 R; 1 N; 0 D; 0 B; 0 C; 2 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 1 I; 2 L; 2 K; 1 M; 1 F; 0 P; 1 S; 1 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:34-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 QVLNIQMRRTLHKAFKGS
 10 X X

12. US-08-249-182-3 (1-5)
 R39830 E1 peptide RV-EP6, residues 106-125.

ID R39830 standard; peptide; 20 AA.
 AC R39830;
 DT 19-JAN-1994 (first entry)
 DE E1 peptide RV-EP6, residues 106-125.
 KW Rubella virus; RV; E1 protein; antibody; mammal; vaccine; rubella;
 KW neutralising; cell mediated; immune response; mumps; measles.
 OS Synthetic.
 PN W09314206-A.
 PD 22-JUL-1993.

PR 20-JAN-1992; GB-001139.
 PA (CONN-) CONNAUGHT LAB LTD.
 PI Chong P, Gillam S, Tingle A;
 DR WPI; 93-243221/30.
 PT Synthetic peptide(s) having at least one antigenic determinant of
 PT a rubella virus protein - useful for producing vaccine, and also
 PT to detect associated antibodies to treat associated auto-immune
 PT disorders, etc.
 PS Table 1; Page 41; 68pp; English.
 CC The sequences given in R39825-52 represent rubella virus (RV) E1
 CC protein peptide fragments. These peptides are capable of eliciting
 CC high titres of antibodies against RV in mammals. They may be used
 CC in vaccines to elicit neutralising antibodies and a cell mediated
 CC immune response against RV. They may be used as one component of a
 CC multivalent vaccine, pref. one providing protection against rubella,
 CC mumps and measles.
 SQ Sequence 20 AA;
 SQ 2 A; 0 R; 0 N; 0 D; 0 B; 1 C; 1 Q; 2 E; 0 Z; 2 G; 2 H;
 SQ 0 I; 0 L; 1 K; 0 M; 1 F; 2 P; 1 S; 1 T; 0 W; 3 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:19-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 GSYVKQYHPTACEVEPAFGH
 10 X 20

13. US-08-249-182-3 (1-5)

R26160 Tuberculosis antibody production peptide #1.

ID R26160 standard; peptide; 20 AA.
 AC R26160;
 DT 02-FEB-1993 (first entry)
 DE Tuberculosis antibody production peptide #1.
 KW 85-C; Mycobacterium tuberculosis; immunise; antibody; vaccine;
 KW carrier molecule; antisera.
 OS Synthetic.
 PN EP-499003-A.
 PD 19-AUG-1992.
 PF 14-FEB-1991; 400388.
 PR 14-FEB-1991; EP-400388.
 PA (INNO-) INNOGENETICS NV SA.
 PI Content J, De Bruyn J, De Wit L;
 DR WPI; 92-277793/34.
 PT Recombinant peptide(s) and their nucleic acids - for diagnosing
 PT tuberculosis and as a vaccine against tuberculosis
 PS Disclosure; Page 11; 48pp; English.
 CC The sequences given in R26160-66 are peptides which were used in the
 CC scope of the invention to raise antibodies against tuberculosis.
 CC They correspond to regions of the 85-C antigen containing region of
 CC Mycobacterium tuberculosis and can be used in the production of
 CC vaccines for immunisation against tuberculosis. The peptides may
 CC be used to raise antisera and in this case would be synthesised with an
 CC additional cysteine residue, pref. attached to the amino terminal.
 CC This facilitates coupling of the peptide to a carrier molecule which
 CC is necessary to render the peptide immunogenic.
 SQ Sequence 20 AA;
 SQ 2 A; 1 R; 2 N; 4 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 2 G; 0 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 1 T; 1 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:21-PDT using FindSeq

WR1; 73-4003347/31.
 PT Hybrid monoclonal antibody - used for prepn. of thrombolytic drug
 PT having increased thrombolytic activity and specificity and
 PT reduced reactivity to fibrinogen
 PS Example 1; Page 14; 38pp; Japanese.
 CC Human fibrin beta-chain peptides A and B were synthesised and coupled
 CC to BSA for injection into mice. The peptides were used to raise
 CC antibodies to human fibrin. Monoclonal antibodies specific for fibrin
 CC are used in the production of bispecific monoclonal antibodies
 CC which also recognise truncated tPA mutants lacking the finger, EGF and
 CC Kringle 1 domains.
 SQ Sequence 8 AA:
 SQ 0 A; 1 R; 0 N; 1 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 0 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 0 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:45-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 2.91
 Residue Identity = 60% Matches = 3 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||
 GHRPLDKC
 X X
 > 0 <
 0| 0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4
 Results file us-08-249-182-3.res made by on Wed 21 Sep 94 12:06:14-PDT.

Query sequence being compared:US-08-249-182-3 (1-5)
 Number of sequences searched: 5543
 Number of scores above cutoff: 1885
 Results of the initial comparison of US-08-249-182-3 (1-5) with:
 File : /home/shears/loring/lorin*.pep

10000-
 -
 N -
 U 5000-
 M *
 B -
 E -
 R - *
 -
 O -
 F 1000-
 -
 S -
 E 500-
 Q -
 U -
 E -
 N -
 C -
 E -
 S 100- *
 -
 -
 50-

111
ERKDVY
X X

14. US-08-249-182-2 (1-6)

P71314 Sequence of fibrin immunogen for the prepn. of mon

ID P71314 standard; peptide; 7 AA.
AC P71314;
DT 19-JUN-1991 (first entry)
DE Sequence of fibrin immunogen for the prepn. of monoclonal antibodies
DE (MAbs).
KW Fibrin-specific monoclonal antibody; screening.
FH Key Location/Qualifiers
FT Misc_difference 7
FT /label= Lys-OH
PN W08706263-A.
PD 22-OCT-1987.
PF 14-APR-1987; U00862.
PR 14-APR-1986; US-851514.
PA (GEHO-) GEN HOSPITAL CORP.
PA (GENO-) GEN HOSPITAL CORP.
PI Matsueda GR, Haber E;
DR WPI; 87-306855/43.
PT Screening of fibrin-specific monoclonal antibodies - by contact
PT with immobilised crosslinked fibrin clot and screening with
PT detectable labelling step
PS Disclosure; Page 7; 41pp; English.
CC The MAbs are specific to fibrin without fibrinogen cross-reactivity.
CC They have increased binding to in vitro and in vivo thrombi. The
CC MAbs can be used in immunoassays for fibrin in the presence of
CC fibrinogen or other proteins. They can be used as immunoaffinity
CC ligands for the purification of fibrin.
SQ Sequence 7 AA;
SQ 0 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 1 G; 1 H;
SQ 0 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 0 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:13-PDT using FindSeq

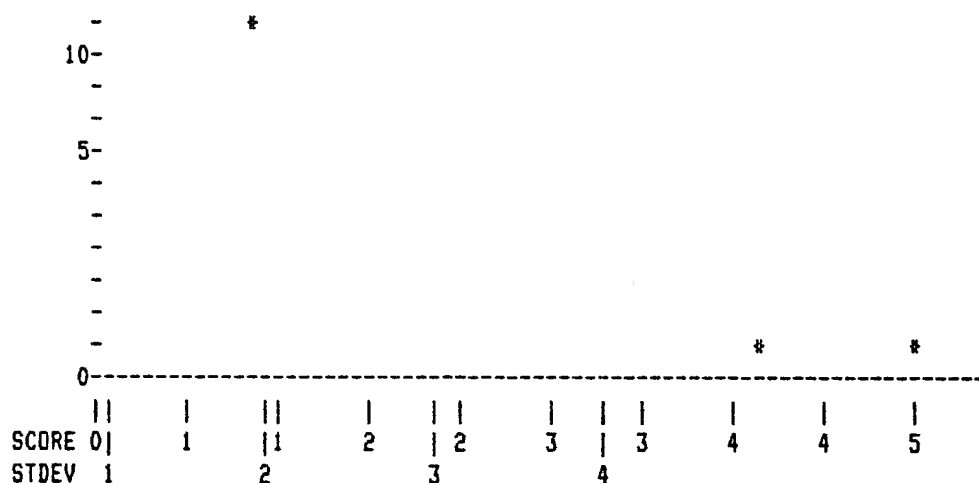
Initial Score = 3 Optimized Score = 3 Significance = 2.91
Residue Identity = 75% Matches = 3 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
111
GHRPLDK
X X

15. US-08-249-182-2 (1-6)

R44837 Human fibrin beta-chain N-terminal peptide A.

ID R44837 standard; peptide; 8 AA.
AC R44837;
DT 20-JUN-1994 (first entry)
DE Human fibrin beta-chain N-terminal peptide A.
KW Tissue plasminogen activator; t-PA; mutein; fibrin; antigen;
KW anti-fibrin; monoclonal antibody; hybridoma; thrombolysis
KW antithrombotic agent; bispecific antibody.
OS Synthetic.
PN J05304992-A.
PD 19-NOV-1993.
PF 17-JUN-1992; 158301.
PR 20-JUN-1991; JP-148936.
PA (TAKE) TAKEDA CHEM IND LTD.



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	0.98
Times:	CPU	Total Elapsed	
	00:00:08.90	00:00:08.00	
Number of residues:	482836		
Number of sequences searched:	5543		
Number of scores above cutoff:	1885		

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37445	Autotaxin peptide ATX 20.	5	5	5	5.12	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
2. R25466	Endoglucanase #2.	376	4	4	4.09	0
**** 3 standard deviations above mean ****						
3. R43439	Rg/SSA epitope 109.	8	3	3	3.07	0

4. R43373	Le/SSb epitope 144.	8	3	3	3.07	0
5. R43440	Ro/SSA epitope 111.	9	3	3	3.07	0
6. R43437	Ro/SSA epitope 105.	10	3	3	3.07	0
7. R07105	Hepatitis B Virus PreS1 epito	11	3	3	3.07	0
8. R43438	Ro/SSA epitope 106.	12	3	3	3.07	0
9. R43392	La/SSb epitope 139.	13	3	3	3.07	0
10. P91772	Mas oncogene angiotensin rec	16	3	3	3.07	0
11. R43391	La/SSb epitope 136.	18	3	3	3.07	0
12. R39830	E1 peptide RV-EP6, residues 1	20	3	3	3.07	0
13. R26160	Tuberculosis antibody product	20	3	3	3.07	0
14. R41297	Peptide fragment F10-2.	21	3	3	3.07	0
15. R34509	Peptide K1 recognised by MAb,	21	3	3	3.07	0
16. R39880	Lipopeptide TPRV-EP6.	23	3	3	3.07	0
17. P20021	Sequence of a foot and mouth	67	3	3	3.07	0
18. R24710	Sequence of a chimeric anti-h	112	3	3	3.07	0
19. P83210	Sequence of the CH2 domain (r	116	3	3	3.07	0
20. R25412	Heavy chain variable domain o	118	3	3	3.07	0
21. R31022	Grass pollen allergen KBG 7.2	131	3	3	3.07	0
22. R24712	Sequence encoded by the genom	132	3	3	3.07	0
23. R27565	Part of lamB lambda receptor	134	3	3	3.07	0
24. A48491	twitching motility protein P	135	3	3	3.07	0
25. A49047	IgM monoclonal striational a	136	3	3	3.07	0
26. R13304	CFTR 556 del A.	151	3	3	3.07	0
27. A30938	myosin regulatory light chai	161	3	3	3.07	0
28. R23867	Pre-S gene region translation	174	3	3	3.07	0
29. R23871	Pre-S gene region translation	174	3	3	3.07	0
30. R23870	Pre-S gene region translation	174	3	3	3.07	0
31. R23869	Pre-S gene region translation	174	3	3	3.07	0
32. R23868	Pre-S gene region translation	174	3	3	3.07	0
33. P90501	ND28 deriv. of G-CSF.	174	3	3	3.07	0
34. R39392	Truncated tissue factor.	218	3	3	3.07	0
35. R43675	Single chain polypeptide with	225	3	3	3.07	0
36. R06478	TRY40.	225	3	3	3.07	0
37. R05710	TRY40.	225	3	3	3.07	0
38. R24811	Sequence encoded by the chim	239	3	3	3.07	0
39. R44510	Type I iodothyronine 5' deiod	249	3	3	3.07	0
40. R14258	gp75 peptide and fragments.	249	3	3	3.07	0

1. US-08-249-182-3 (1-5)

R37445 Autotaxin peptide ATX 20.

ID R37445 standard; peptide; 5 AA.
AC R37445;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 20.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 20. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin

CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 5 AA;
SQ 1 A; 0 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 1 K; 0 M; 1 F; 1 P; 0 S; 0 T; 0 W; 1 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 5.12
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
YPAFK
X X

2. US-08-249-182-3 (1-5)
R25466 Endoglucanase #2.

ID R25466 standard; Protein; 376 AA.
AC R25466;
DT 07-JAN-1993 (first entry)
DE Endoglucanase #2.
KW CMC-endoase; 43 kD cellulase; monoclonal antibody.
OS Humicola insolens.
PN EP-495257-A.
PD 22-JUL-1992.
PF 06-NOV-1991; 202879.
PR 16-JAN-1991; EP-870006.
PR 06-NOV-1991; EP-202880.
PR 06-NOV-1991; EP-202879.
PA (PRDC) PROCTER & GAMBLE CO.
PI Baeck AC, Busch A, Ceulemans RAA;
DR WPI; 92-243163/30.
DR N-PSDB; 026382.
PT Compact, granular detergent compsns. - contain high activity
PT cellulase and softening clay to provide synergistic effect in
PT softening performance
PS Disclosure; Page 23-25; 29pp; English.
CC The sequences given in R25464 and R25466 are endoglucanases which
CC are immunoreactive with a monoclonal antibody raised against a
CC partially purified 43 kD cellulase derived from Humicola insolens.
CC These endoglucanases exhibit a CMC-endoase activity of at least
CC about 50, pref. at least about 60, inparticular at least about 90 CMC-
CC endoase units per µg of total protein. These endoglucanases have
CC molecular weight of approx. 43 kD.
SQ Sequence 376 AA;
SQ 46 A; 7 R; 17 N; 22 D; 0 B; 19 C; 15 Q; 10 E; 0 Z; 34 G; 3 H;
SQ 6 I; 15 L; 33 K; 4 M; 11 F; 29 P; 38 S; 29 T; 6 W; 13 Y; 19 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:18-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 4.09
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
DSYPELLKDGCHWRFDWFENADNPDTFEQVQCPKALLDISGCKRDDSSFPAPFKVDTSASKPQPSSSAKKT
180 190 200 210 220 230 240

TSAAAAAQPQKTKDSAPVVQKSSTKPAAGPEPT
250 260 270 280

3. US-08-249-182-3 (1-5)

R43439 Ro/SSA epitope 109.

ID R43439 standard; peptide; 8 AA.
AC R43439;
DT 12-MAY-1994 (first entry)
DE Ro/SSA epitope 109.
KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
OS Homo sapiens.
PN W09321223-A.
PD 28-OCT-1993.
PF 13-APR-1993; U03484.
PR 13-APR-1992; US-867819.
PA (OKLA) UNIV OKLAHOMA STATE.
PI Harley JB;
DR WPI; 93-351658/44.
PT New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus erythematosus
PS Claim 1; Page 30; 43pp; English.
CC The sequences given in R43391-562 are linear epitopes which are
CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polyuridine
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human
CC autoantibody and may therefore be used as vaccines.
SQ Sequence 8 AA;
SQ 3 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 2 K; 0 M; 1 F; 0 P; 0 S; 0 T; 0 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:35-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 60% Matches = 3 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
KGAAFKAV
X X

4. US-08-249-182-3 (1-5)

R43393 La/SSb epitope 144.

ID R43393 standard; peptide; 8 AA.
AC R43393;

DE La/SSB epitope 144.
KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
OS Homo sapiens.
PN W09321223-A.
PD 28-OCT-1993.
PF 13-APR-1993; U03484.
PR 13-APR-1992; US-867819.
PA (OKLA) UNIV OKLAHOMA STATE.
PI Harley JB;
DR WPI; 93-351658/44.
PT New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus erythematosus
PS Claim 1; Page 30; 43pp; English.
CC The sequences given in R43391-562 are linear epitopes which are
CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polyuridine
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human
CC autoantibody and may therefore be used as vaccines.
SQ Sequence 8 AA;
SQ 1 A; 1 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 1 H;
SQ 0 I; 1 L; 2 K; 0 M; 1 F; 0 P; 0 S; 1 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:34-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 60% Matches = 3 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
RTLHKAFK
X X

5. US-08-249-182-3 (1-5)

R43440 Ro/SSA epitope 111.

ID R43440 standard; peptide; 9 AA.
AC R43440;
DT 12-MAY-1994 (first entry)
DE Ro/SSA epitope 111.
KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
OS Homo sapiens.

US Homo sapiens.
 PN W09321223-A.
 PD 28-OCT-1993.
 PF 13-APR-1993; U03484.
 PR 13-APR-1992; US-867819.
 PA (OKLA) UNIV OKLAHOMA STATE.
 PI Harley JB;
 DR WPI; 93-351658/44.
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 PS Claim 1; Page 30; 43pp; English.
 CC The sequences given in R43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 SQ Sequence 9 AA;
 SQ 2 A; 0 R; 0 N; 0 D; 0 B; 1 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 1 K; 0 M; 1 F; 0 P; 1 S; 0 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:35-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 100% Matches = 3 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||
 AFKAVSEVC
 X X

6. US-08-249-182-3 (1-5)
 R43437 Ro/SSA epitope 105.

ID R43437 standard; peptide; 10 AA.
 AC R43437;
 DT 12-MAY-1994 (first entry)
 DE Ro/SSA epitope 105.
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 PN W09321223-A.
 PD 28-OCT-1993.
 PF 13-APR-1993; U03484.
 PR 13-APR-1992; US-867819.
 PA (OKLA) UNIV OKLAHOMA STATE.

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 60% Matches = 3 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
DGLRAQDDYNGWDINTPAFE
10 X 20

14. US-08-249-182-3 (1-5)

R41297 Peptide fragment F10-2.

ID R41297 standard; Protein; 21 AA.
AC R41297;
DT 20-APR-1994 (first entry)
DE Peptide fragment F10-2.
KW Pyruvate dehydrogenase complex; PDC;
KW oxoglutarate dehydrogenase complex; OGDC;
KW monoclonal antibody; E2 enzyme; antigen.
PN J05227984-A.
PD 07-SEP-1993.
PF 25-FEB-1992; 037749.
PR 25-FEB-1992; JP-037749.
PA (SUMO) SUMITOMO CHEM CO LTD.
PA (SUMU) SUMITOMO SEIYAKU KK.
DR WPI; 93-316617/40.
PT Monoclonal antibody prepd. by culturing new hybridoma - used for
PT determn. of pyruvate dehydrogenase complex and oxo-glutarate
PT dehydrogenase complex in disease research
PS Disclosure; Page 9; 11pp; English.
CC Pyruvate dehydrogenase complex (PDC) and oxoglutarate dehydrogenase
CC complex (OGDC) may be detected using a monoclonal antibody specifically
CC recognising E2 enzyme contg. in PDC and PGDC.
SQ Sequence 21 AA;
SQ 2 A; 1 R; 1 N; 2 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 1 G; 0 H;
SQ 1 I; 0 L; 0 K; 0 M; 1 F; 1 P; 1 S; 0 T; 0 W; 0 Y; 2 V;
SQ 6 Others;
CC Retrieved by shears on Wed 21 Sep 94 11:59:31-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
Residue Identity = 60% Matches = 3 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||
NDVIXXQXPFAEXXSXGDVR
10 X 20

15. US-08-249-182-3 (1-5)

R34509 Peptide K1 recognised by MAb, for assaying LACI.

ID R34509 standard; peptide; 21 AA.
AC R34509;
DT 10-AUG-1993 (first entry)
DE Peptide K1 recognised by MAb, for assaying LACI.
KW LACI; lipoprotein associated coagulation inhibitor; EPI;
KW extrinsic pathway inhibitor; blood clotting; TFPI;
KW tissue factor pathway inhibitor; blood clotting; blood coagulation;
KW immunological assay; ELISA; testing; detection; diagnosis;
KW coronary thrombosis; pulmonary thrombosis; thrombotic disease;
KW microthrombi; monoclonal antibody; MAb.

PN EP-539975-A.
 PD 05-MAY-1993.
 PF 29-OCT-1992; 118497.
 PR 31-OCT-1991; JP-311442.
 PR 29-NOV-1991; JP-339560.
 PA (TEIJ) TEIJIN LTD.
 PI Ichikawa Y, Koike Y, Suzuki K;
 DR WPI; 93-145355/18.
 PT Assay method for lipoprotein-associated coagulation inhibitor -
 PT using monoclonal antibodies; used for diagnosis of thrombosis
 PS Claim 1; Page 12; 17pp; English.
 CC This sequence (K1) is recognised by a monoclonal antibody used in an
 CC immunological assay for free lipoprotein associated coagulation
 CC inhibitor (LACI). The MAb is used together with another MAb
 CC recognising peptide K3 (R34508).
 SQ Sequence 21 AA;
 SQ 3 A; 1 R; 1 N; 2 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 2 I; 0 L; 3 K; 1 M; 5 F; 1 P; 0 S; 0 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:03-PDT using FindSeq

Initial Score = 3 Optimized Score = 3 Significance = 3.07
 Residue Identity = 100% Matches = 3 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

      X X
    YPAFK
      |||
      AFKADDGPCKAIMKRFFFNIF
      X X      10      20
> D <
D| |D IntelliGenetics
> D <
  
```

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-08-249-182-4.res made by on Wed 21 Sep 94 12:08:17-PDT.

Query sequence being compared:US-08-249-182-4 (1-5)
 Number of sequences searched: 5543
 Number of scores above cutoff: 2112

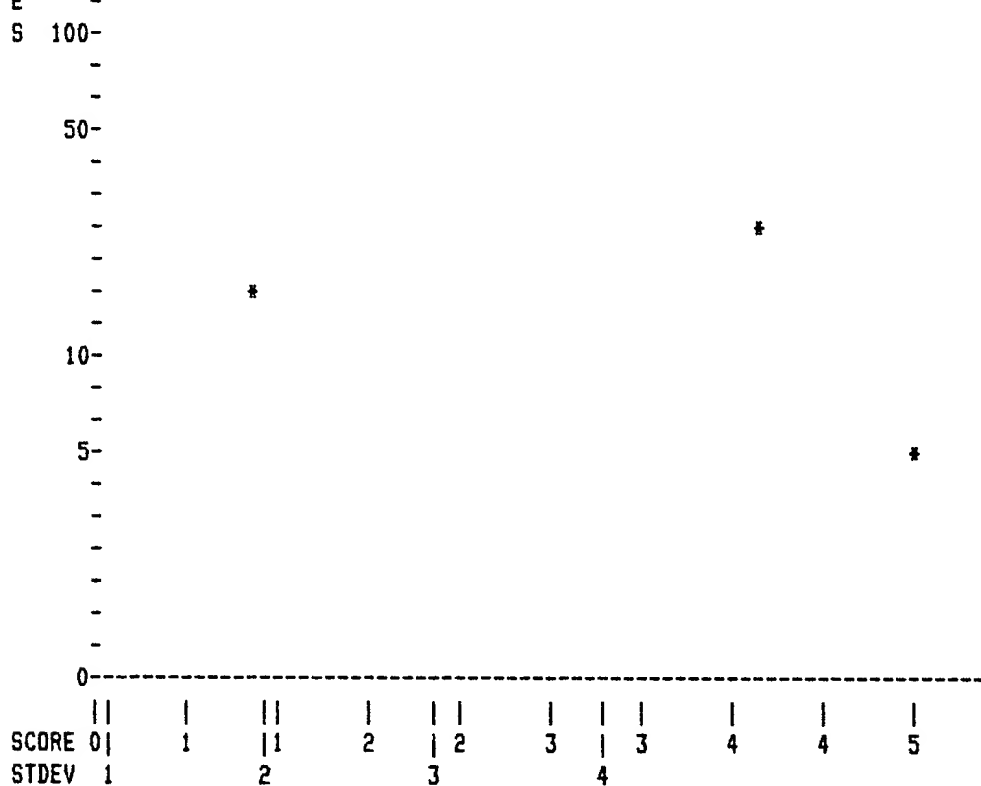
Results of the initial comparison of US-08-249-182-4 (1-5) with:
 File : /home/shears/loring/loring*.pep

```

10000-
-
N -
U 5000-
M *
B -
E -
R -
-
O -
F 1000-
-
S -
E 500-
Q -
U -
E -
N -
C -
  
```

*

*



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.07

Times:	CPU	Total Elapsed
	00:00:08.90	00:00:09.00

Number of residues:	482836
Number of sequences searched:	5543
Number of scores above cutoff:	2112

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37446	Autotaxin peptide ATX 24.	5	5	5	4.67	0

4 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
---------------	-------------	--------	-------------	------------	------	-------

2. R26150	ASB7 gH-2 antibody grafted he	146	5	5	4.67	0
3. R20793	CDR-grafted, humanised heavy	146	5	5	4.67	0
4. R41469	MAB 25D2 humanised heavy chai	140	5	5	4.67	0
5. R40179	Humanised antibody CMX5-3 hea	135	5	5	4.67	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 3 standard deviations above mean ****						
6. R41113	HCV peptide XIV or HCV8 (aa 1	22	4	4	3.74	0
7. R41186	HCV NS4 protein HCV7/8.	26	4	4	3.74	0
8. R40053	Hib OMP P1 peptide HIBP1-1 (1	30	4	4	3.74	0
9. R13186	Peptide VII immunoreactive wi	38	4	4	3.74	0
10. R06408	HTLV-1 corresponding peptide	38	4	4	3.74	0
11. R13185	Peptide (V) immunoreactive wi	40	4	4	3.74	0
12. R13591	HTLV-I env precursor epitope	40	4	4	3.74	0
13. R33871	Polypeptide p1684 comprising	67	4	4	3.74	0
14. R33872	Polypeptide p1689 comprising	117	4	4	3.74	0
15. R11717	ENV93/HTLV-1-II fusion protei	217	4	4	3.74	0
16. R39336	scFv fragment encoded by pLIS	277	4	4	3.74	0
17. R11718	ENV93/HTLV-1-II+I fusion prot	344	4	4	3.74	0
18. R23999	Open reading frame of the hep	363	4	4	3.74	0
19. P70416	Polypeptide with IgE binding	557	4	4	3.74	0
20. R26207	Human serum albumin.	585	4	4	3.74	0
21. P70417	Polypeptide with IgE binding	775	4	4	3.74	0
22. FRZE_MYXXA	GLIDING MOTILITY REGULATORY P	777	4	4	3.74	0
23. R04574	Derived amino acid sequence o	810	4	4	3.74	0
24. P98500	Partial sequence encoded by m	1091	4	4	3.74	0
25. R28582	HCV amino acid sequence contg	2436	4	4	3.74	0
26. R24440	Composite HCV HC-J1/CDC/CHI p	2894	4	4	3.74	0
27. R31621	Hepatitis C virus (HCV) polyp	3011	4	4	3.74	0
28. P90373	Sequence encoded by human mus	3685	4	4	3.74	0
**** 2 standard deviations above mean ****						
29. R41316	PEP (84-91).	8	3	3	2.80	0
30. R36072	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
31. R36071	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
32. R36070	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
33. R36069	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
34. R36068	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
35. R36067	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
36. R36024	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
37. R36023	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
38. R36022	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
39. R36021	Hepatitis C virus (HCV) epito	8	3	3	2.80	0
40. R36020	Hepatitis C virus (HCV) epito	8	3	3	2.80	0

1. US-08-249-182-4 (1-5)

R37446 Autotaxin peptide ATX 24.

ID R37446 standard; peptide; 5 AA.
AC R37446;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 24.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 24. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 5 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 1 S; 0 T; 0 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.67
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 QAEVS
 X X

2. US-08-249-182-4 (1-5)

R26150 A5B7 gH-2 antibody grafted heavy chain.

ID R26150 standard; Protein; 146 AA.
 AC R26150;
 DT 03-FEB-1993 (first entry)
 DE A5B7 gH-2 antibody grafted heavy chain.
 KW humanised antibody; chimaeric; carcino-embryonic antigen; therapy;
 KW diagnosis; carcinomas; CDR; IgG; human; murine; ss.
 OS Chimaeric.
 FH Key Location/Qualifiers
 FT Region 26..35
 FT /note= "grafted murine CDR1"
 FT Region 50..65
 FT /note= "grafted murine CDR2"
 FT Region 95-102
 FT /note= "grafted murine CDR3"
 PN W09201059-A.
 PD 23-JAN-1992.
 PF 05-JUL-1991; G01108.
 PR 05-JUL-1990; GB-014932.
 PR 21-DEC-1990; WD-G02017.
 PR 05-JUL-1991; WD-G01108.
 PA (CELL-) CELLTECH LTD.
 PI Adair JR, Bodmer MW, Mountain A, Owens RJ;
 DR WPI; 92-284316/34.
 DR N-PSDB; 027354.
 PT Humanised antibody molecules - comprising murine and human regions,
 PT specific for carcino-embryonic antigen, useful for diagnosis and
 PT therapy
 PS Example 4; Figure 10; 71pp; English.
 CC This sequence is CDR-grafted A5B7 human antibody having
 CC murine CDRs at amino acids 26-35 (CDR1), 50-65 (CDR2), and 95-102
 CC (CDR3) and additional murine framework residues at 1, 48, 49, 72,
 CC 73, 76, and 93. The LAY framework was chosen when making the coding
 CC construct (027354) as it shows the highest homology to A5B7. The
 CC antibody has specificity for carcinoembryonic antigen, produced by

CC certain carcinomas.
 SQ Sequence 146 AA;
 SQ 7 A; 7 R; 4 N; 4 D; 0 B; 2 C; 6 Q; 6 E; 0 Z; 18 G; 1 H;
 SQ 3 I; 14 L; 6 K; 3 M; 9 F; 3 P; 16 S; 13 T; 5 W; 9 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:22-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.67
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 |||||

QAPGKGLEWLGFIGNKANGYITEYSASVKGRFTISRDKSKSTLYLQMNGLQAEVSAIYYCTDRGLRFYFDY
 60 70 80 90 100 110 X 120

WGQGT LVT VSSASTKGP
 130 140

3. US-08-249-182-4 (1-5)

R20793 CDR-grafted, humanised heavy chain gH1.

ID R20793 standard; Protein; 146 AA.
 AC R20793;
 DT 19-MAY-1992 (first entry)
 DE CDR-grafted, humanised heavy chain gH1.
 KW murine monoclonal antibody; MAb; A5B7; humanised antibody; CEA;
 KW complementarity determining region.
 OS Homo sapiens.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label= signal
 FT Protein 20..146
 FT /label= VH
 FT /note= "human LAY framework with A5B7 CDRs"
 FT Region 45..54
 FT /label= CDR1
 FT /note= "murine residues"
 FT Region 69..87
 FT /label= CDR2
 FT /note= "murine residues"
 FT Region 120..129
 FT /label= CDR3
 FT /note= "murine residues"
 FT Misc_difference 20
 FT /note= "murine residue"
 FT Misc_difference 67..68
 FT /note= "murine residues"
 FT Misc_difference 94..95
 FT /note= "murine residues"
 FT Misc_difference 98
 FT /note= "murine residue"
 FT Misc_difference 118
 FT /note= "murine residue"
 PN WD9201059-A.
 PD 23-JAN-1992.
 PF 05-JUL-1991; G01108.
 PR 05-JUL-1990; GB-014932.
 PR 21-DEC-1990; WD-G02017.
 PR 05-JUL-1991; WD-G01108.
 PA (CELL-) CELLTECH LTD.
 PI Adair JR, Bodmer MW, Mountain A, Owens RJ;
 DR WPI; 92-056874/07.

PT New CDR-grafted anti carcinoembryonic antigen antibodies - useful
 PT in therapy and diagnosis of carcinoma
 PS Claim 14; Fig 10; 70pp; English.
 CC This heavy chain sequence comprises a human framework (i.e. the LAY
 CC region) which contains murine sequences (from the murine anti-CEA
 CC A5B7 MAb) in the CDRs and at other positions predicted to be
 CC important for antigen-binding and at which human and A5B7 sequences
 CC differ. (See Q20984 for A5B7 heavy chain coding sequence).
 SQ Sequence 146 AA;
 SQ 7 A; 7 R; 4 N; 4 D; 0 B; 2 C; 6 Q; 6 E; 0 Z; 18 G; 1 H;
 SQ 3 I; 14 L; 6 K; 3 M; 9 F; 3 P; 16 S; 13 T; 5 W; 9 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:50-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.67
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 I I I I
 QAPGKGLEWLGFIGNKANGYTTSEYASVKGRTISRDKSKSTLYLQMNGLQAEVSAIYYCTDRGLRFYFDY
 60 70 80 90 100 110 X 120
 MGQGLTVTVSSASTKGP
 130 140

4. US-08-249-182-4 (1-5)

R41469 MAb 25D2 humanised heavy chain.

ID R41469 standard; Protein; 140 AA.
 AC R41469;
 DT 03-MAR-1994 (first entry)
 DE MAb 25D2 humanised heavy chain.
 KW Heavy; VH; light; VL; chain; variable region; antihuman; interleukin-4;
 KW IL-4; monoclonal antibody; MAb; 25D2; single chain binding protein;
 KW complementarity determining region; CDR; humanised; Fv region; BABS;
 KW antagonist; polymerase chain reaction; PCR; primer; amplify; gamma4;
 KW pSV.SPORT.
 OS Rattus rattus.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /note= "Leader sequence"
 FT Protein 20..140
 FT /note= "25D2 H chain"
 PN W09317106-A.
 PD 02-SEP-1993.
 PF 18-FEB-1993; U01301.
 PR 19-FEB-1992; US-841659.
 PA (SCHE) SCHERING CORP.
 PI Abrams JS, Dalie B, Le HV, Miller K, Murgolo NJ;
 PI Nguyen H, Pearce M, Tindall S, Zavodny PJ;
 DR WPI; 93-288412/36.
 DR N-PSDB; Q48350.
 PT Monoclonal antibodies against human interleukin-4 corresp. DNA
 PT and CDRs - are useful for detection of interleukin-4 and treatment
 PT of related diseases
 PS Example 9; Page 89-90; 114pp; English.
 CC This sequence represents the humanised heavy (H) chain of the antihuman
 CC interleukin-4 (IL-4) monoclonal antibody (MAb) 25D2. The 25D2 H
 CC chain coding region was cloned in three fragments using the primers
 CC given in Q48351-60. The amplified fragments were designed to
 CC contain silent restriction sites, however several codons had to
 CC be changed to incorporate further restriction sites. The primers

CC given in 44551-55 were used to amplify the entire H chain variable
 CC region (VH) of an unrelated humanised antibody. The amplified fragments
 CC were then cloned into pSV.Sport which already contained the 25D2 H
 CC chain fragments. The primers given in 048367-72 were used in
 CC further manipulations to amplify a human gamma4 constant region cDNA
 CC which was used to replace the genomic DNA. The humanised MAb is an
 CC IL-4 antagonist. It may be used in a pharmaceutical composition for
 CC detecting, measuring and immunopurifying human IL-4 and blocking IL-4
 CC activity in IL-4-related diseases.
 SQ Sequence 140 AA;
 SQ 9 A; 7 R; 4 N; 5 D; 0 B; 3 C; 6 Q; 4 E; 0 Z; 16 G; 2 H;
 SQ 6 I; 12 L; 2 K; 3 M; 7 F; 4 P; 18 S; 8 T; 5 W; 9 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:27-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.67
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 0AEVS
 |||||
 VR0APGKGLEWVASISISGDNITYYPDSVRGRFTISRND SKNTLYLQMNGL0AEVSAIYYCARDPYFSGHYF
 60 70 80 90 100 X 110 120
 DFWG0GTLVTVSS
 130 140

5. US-08-249-182-4 (1-5)

R40179 Humanised antibody CMX5-3 heavy chain variable reg

ID R40179 standard; Protein; 135 AA.
 AC R40179;
 DT 14-FEB-1994 (first entry)
 DE Humanised antibody CMX5-3 heavy chain variable region.
 KW Primer; polymerase chain reaction; amplify; PCR; human; kappa; L;
 KW constant region; heavy; H; chain; pUC19; humanised; antibody;
 KW light; REI; VL3 fragment; CMX5-1; CMX5-3.
 DS Synthetic.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /note= "Secretory leader peptide"
 FT Protein 20..135
 FT /note= "CMX5-3 heavy chain variable region"
 PN W09316184-A.
 PD 19-AUG-1993.
 PF 04-FEB-1993; U00759.
 PR 06-FEB-1992; US-832842.
 PA (SCHE) SCHERING CORP.
 PI Abrams JS, Chou C, Jenh C, Murgolo NJ, Petro ME;
 PI Silver JE, Tindall S, Windsor WT, Zavodny PJ;
 DR WPI; 93-272888/34.
 PT Humanised monoclonal antibody - comprises variable animal region
 PT and constant human region, binds to human interleukin-5
 PS Example; Page 91-92; 118pp; English.
 CC The sequences given in R40179-80 represent the variable regions of
 CC the heavy and light chains of the humanised antibody CMX5-3
 CC respectively. These sequences were based on the humanised antibody
 CC CMX5-1. These sequences were generated using the primer sequences
 CC given in 048068-71. These primers were based on sequences derived
 CC from antibody JES1-39D10 and human LAY VH framework sequences. The
 CC amplification products were used to replace the VH1 and VH3 fragments
 CC of CMX5-1 H chain cDNA in pSV.Sport (see also R40175).
 SQ Sequence 135 AA;
 SQ 9 A; 5 R; 7 N; 4 D; 0 B; 3 C; 6 Q; 5 E; 0 Z; 15 G; 1 H;
 SQ 7 I; 15 L; 3 K; 2 M; 4 F; 2 P; 17 S; 8 T; 5 W; 7 Y; 10 V;

CC useful for the detection of antibodies to HCV, and/or HIV,
CC and/or HTLV-I or II.
SQ Sequence 22 AA;
SQ 5 A; 1 R; 0 N; 0 D; 0 B; 0 C; 3 Q; 1 E; 0 Z; 1 G; 0 H;
SQ 1 I; 3 L; 1 K; 0 M; 0 F; 1 P; 1 S; 1 T; 0 W; 0 Y; 1 V;
SQ 2 Others;
CC Retrieved by shears on Wed 21 Sep 94 11:59:30-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||
XGKALGLLQTASRQAEVIAPAX
10 X 20

7. US-08-249-182-4 (1-5)

R41186 HCV NS4 protein HCV7/8.

ID R41186 standard; peptide; 26 AA.
AC R41186;
DT 22-MAR-1994 (first entry)
DE HCV NS4 protein HCV7/8.
KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
OS Synthetic.
FH Key Location/Qualifiers
FT Region 10..16
FT /label= epitope_4
PN WD9318054-A.
PD 16-SEP-1993.
PF 08-MAR-1993; E00517.
PR 06-MAR-1992; EP-400598.
PA (INNO-) INNOGENETICS NV SA.
PI De LEYS R;
DR WPI; 93-303397/38.
PT New biotinylated peptide(s) corresp. to immuno-dominant
PT epitope(s) - with increased antigenicity, useful in antibodies
PT detection and vaccines against hepatitis C, HIV and HTLV
PS Disclosure; Page 79; 133pp; English.
CC Peptide compsns. comprise at least one and pref. a combination of
CC two, three, four or more biotinylated peptides chosen from the
CC sequences given in R41058-R41166.
CC The peptides may be hybrids consisting of combinations of the core
CC epitopes of the HCV core (R41171-R41180), HCV NS4 (R41181-R41186) or
CC the HCV NS5 (R41187-R41193) region separated by Gly and/or Ser residues.
CC Pref. hybrid peptides are given in R41161-R41163.
CC The peptides represent immunologically important regions of viral
CC proteins and are prep'd. by solid phase peptide synthesis. The compsns.
CC are useful for the detection of antibodies to HCV, and/or HIV,
CC and/or HTLV-I or II.
SQ Sequence 26 AA;
SQ 6 A; 1 R; 0 N; 0 D; 0 B; 0 C; 4 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 4 L; 2 K; 0 M; 1 F; 1 P; 1 S; 1 T; 0 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:29-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS

Initial Score = 5 Optimized Score = 5 Significance = 4.67
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                |||||
WIRQAPGKGLEWVALIWSNGD TDYNSAIKSRFTISRND SKNTLYLQMNGLQAEVSAIYFCAREYYGYFDYWG
   60       70       80       90      100    X 110      120

QGTLVTVSS
130

```

6. US-08-249-182-4 (1-5)

R41113 HCV peptide XIV or HCV8 (aa 1730-1749).

ID R41113 standard; peptide; 22 AA.
 AC R41113;
 DT 22-MAR-1994 (first entry)
 DE HCV peptide XIV or HCV8 (aa 1730-1749).
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;
 FT A (optional)= one or more amino acids, NH2 or
 FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH2, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 22
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN WD9318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT detection and vaccines against hepatitis C, HIV and HTLV
 PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are

8. US-08-249-182-4 (1-5)

R40053 Hib OMP P1 peptide HIBP1-1 (1-29).

ID R40053 standard; peptide; 30 AA.
AC R40053;
DT 04-FEB-1994 (first entry)
DE Hib OMP P1 peptide HIBP1-1 (1-29).
KW Haemophilus influenzae; type b; Hib; outer membrane protein; P1; P2;
KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
KW immunogen.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 30
FT /note= "May be absent"
PN W09315205-A.
PD 05-AUG-1993.
PF 03-FEB-1993; CA0041.
PR 03-FEB-1992; GB-002219.
PA (CONN-) CONNAUGHT LAB LTD.
PI Chong P, Kandil A, Klein MH, Sia C;
DR WPI; 93-258681/32.
PT Synthetic Haemophilus influenzae conjugate vaccine - comprising
PT T-helper cell determinants and B-cell epitope(s) linked to
PT synthetic oligo:saccharide(s)
PS Table 1; Page 47; 99pp; English.
CC The sequences given in R40053-101 are peptide fragments derived from
CC the Haemophilus influenzae type b (Hib) outer membrane proteins P1,
CC P2 and P6. These peptides may be used in a vaccine against Hib
CC infection and antibodies against these peptides may be used in test
CC kits to detect H. influenzae in a sample. The vaccine may further
CC comprise a immunogenic or immunostimulatory molecule or the peptides
CC may be modified with lipids, or linked to synthetic PRP as synthetic
CC lipoglycopeptide conjugates to produce alternative vaccines.
SQ Sequence 30 AA;
SQ 9 A; 1 R; 1 N; 1 D; 0 B; 1 C; 1 Q; 2 E; 0 Z; 3 G; 0 H;
SQ 1 I; 2 L; 0 K; 0 M; 1 F; 0 P; 3 S; 1 T; 0 W; 1 Y; 2 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:22-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||
AAFQLAEVSTSGLGGRAYAGEAAIADNASVC
X 10 20 30

9. US-08-249-182-4 (1-5)

R13186 Peptide VII immunoreactive with anti-HTLV antibody

ID R13186 standard; Protein; 38 AA.
AC R13186;
DT 09-OCT-1991 (first entry)
DE Peptide VII immunoreactive with anti-HTLV antibodies.
KW human T-cell leukaemia virus; AIDS; ATL; detection;
KW envelope protein gp61; acquired immunodeficiency syndrome.
OS Synthetic.
PN EP-439077-A.
PD 31-JUL-1991.

PAIIPDREVLRYREFDEMEECSSQHLPLYIEQGMMMLAEQFKQKALGLLQTASRQAEVIAPAV
10 20 30 40 50 60 X

14. US-08-249-182-4 (1-5)

R33872 Polypeptide p1689 comprising HCV viral antigen.

ID R33872 standard; peptide; 117 AA.
AC R33872;
DT 19-JUL-1993 (first entry)
DE Polypeptide p1689 comprising HCV viral antigen.
KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
KW p408.
OS Synthetic.
PN W09306247-A.
PD 01-APR-1993.
PF 16-SEP-1992; U07813.
PR 16-SEP-1991; US-760292.
PA (ABB0) ABBOTT LAB.
PI Lesniewski RR, Leung TK;
DR WPI; 93-117563/14.
PT Assay for detecting presence of antibody to hepatitis C viral
PT antigen - by contacting sample with polypeptide contg. at least
PT one epitope of virus antigen
PS Disclosure; Page 13; 63pp; English.
CC The synthetic peptide p1689 represents amino acid residues 1689-1805 of
CC the hepatitis C viral antigen. The peptide may be used in an assay to
CC detect antibodies to HCV and thus to diagnose chronic HCV infection.
CC See also R33861-87.
SQ Sequence 117 AA;
SQ 14 A; 3 R; 3 N; 2 D; 0 B; 1 C; 10 Q; 9 E; 0 Z; 6 G; 2 H;
SQ 7 I; 12 L; 5 K; 5 M; 5 F; 7 P; 8 S; 8 T; 3 W; 3 Y; 4 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:46-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
IIII
PAIIPDREVLRYREFDEMEECSSQHLPLYIEQGMMMLAEQFKQKALGLLQTASRQAEVIAPAVQTNWQKLETFWAK
10 20 30 40 50 X 60 70
HMWNFISGIQVLAGLSTLPGNPAIASLMAFTAA
80 90 100

15. US-08-249-182-4 (1-5)

R11717 ENV93/HTLV-1-II fusion protein.

ID R11717 standard; Protein; 217 AA.
AC R11717;
DT 27-JUN-1991 (first entry)
DE ENV93/HTLV-1-II fusion protein.
KW Human T-cell leukaemia virus; HTLV-1; fusion protein; antibodies.
OS Human T-cell leukaemia virus.
FH Key Location/Qualifiers
FT Region 1..104
FT /label= ENV93 epitope
FT Region 105..217
FT /label= HTLV-1-II epitope
PN EP-424748-A.
PD 02-MAY-1991.

PA (TRIT-) TRITON BIOSCIENCES.
 PI Akita RW, Florine DL, Ralston JS;
 DR WPI; 91-221557/30.
 PT Synthetic peptide(s) and antibodies corresp. to an epitope of
 PT HTLV-I - used in diagnosis, therapy prepn. of vaccines and
 PT prognostic indicators of HTLV-I infection
 PS Disclosure; Page 3; 10pp; English.
 CC The peptide has specific binding affinity for 0.5alpha monoclonal
 CC antibody. It represents an epitopic site on the major HTLV-I
 CC envelope precursor, i.e it corresponds to residues 281-320. The
 CC peptide and its antibody can be used in diagnosing the presence of
 CC HTLV-I associated diseases, as vaccines against HTLV-I infection or
 CC as prognostic indicators after HTLV-I infection.
 CC See also R13077, R13590 and R13591.
 SQ Sequence 40 AA;
 SQ 3 A; 3 R; 0 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 3 I; 5 L; 0 K; 0 M; 1 F; 6 P; 7 S; 1 T; 1 W; 0 Y; 5 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:30-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 GAEVS
 || ||
 IQAIVSSPCHCSLILPPFSLSPVPTLGSRRRAVPVAVWL
 X X 10 20 30 40

13. US-08-249-182-4 (1-5)

R33871 Polypeptide p1684 comprising HCV viral antigen.

ID R33871 standard; peptide; 67 AA.
 AC R33871;
 DT 19-JUL-1993 (first entry)
 DE Polypeptide p1684 comprising HCV viral antigen.
 KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
 KW p408.
 OS Synthetic.
 PN W09306247-A.
 PD 01-APR-1993.
 PF 16-SEP-1992; U07813.
 PR 16-SEP-1991; US-760292.
 PA (ABB0) ABBOTT LAB.
 PI Lesniewski RR, Leung TK;
 DR WPI; 93-117563/14.
 PT Assay for detecting presence of antibody to hepatitis C viral
 PT antigen - by contacting sample with polypeptide contg. at least
 PT one epitope of virus antigen
 PS Disclosure; Page 13; 63pp; English.
 CC The synthetic peptide p1684 represents amino acid residues 1684-1750 of
 CC the hepatitis C viral antigen. The peptide may be used in an assay to
 CC detect antibodies to HCV and thus to diagnose chronic HCV infection.
 CC See also R33861-87.
 SQ Sequence 67 AA;
 SQ 7 A; 4 R; 0 N; 2 D; 0 B; 1 C; 6 Q; 8 E; 0 Z; 4 G; 1 H;
 SQ 4 I; 7 L; 3 K; 3 M; 2 F; 4 P; 3 S; 1 T; 0 W; 2 Y; 5 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:46-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```
      X  X
      QAEVS
      || ||
CFDPQIQAI VSSPCHNSLILPPFSLSPVPTLGSRSRRA
      X 10      20      30
```

11. US-08-249-182-4 (1-5)

R13185 Peptide (V) immunoreactive with anti-HTLV antibody

ID R13185 standard; Protein; 40 AA.
AC R13185;
DT 09-OCT-1991 (first entry)
DE Peptide (V) immunoreactive with anti-HTLV antibodies.
KW human T-cell leukaemia virus; AIDS; ATL; detection;
KW envelope protein gp61; acquired immunodeficiency syndrome.
OS Synthetic.
PN EP-439077-A.
PD 31-JUL-1991.
PF 18-JAN-1991; 100616.
PR 24-JAN-1990; US-469291.
PA (UNBI-) UNITED BIOMEDICAL.
PI Wang CY;
DR WPI; 91-224505/31.
PT Peptide compsns. corresp. to envelope fragments of HTLV-1,2 - for
PT detecting antibodies to these viruses and diagnosing HIV and
PT adult T-cell leukaemia infections
PS Claim 1; Page 17; 27pp; English.
CC This peptide is one of 16 peptides useful for detecting antibodies to
CC HTLV or HIV viruses. The peptides correspond to partial sequences of
CC the HTLV virus designated gp21 and gp64, both part of gp61, which
CC defines the envelope protein of the HTLV-I or HTLV-II virus. The
CC peptides can be amidated at the C-terminal. This particular peptide
CC is used in a composition with at least two of the other peptides of
CC the invention. See also R13184, R13186-R13192 and R13861-6.
SQ Sequence 40 AA:
SQ 3 A; 0 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 0 G; 3 H;
SQ 2 I; 6 L; 0 K; 0 M; 2 F; 7 P; 5 S; 3 T; 1 W; 1 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:31-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```
      X  X
      QAEVS
      || ||
SSTPLLYPSLALPAPHLTLPFNWTHCFDPQIQAI VSSPCH
      10      20      30 X  X 40
```

12. US-08-249-182-4 (1-5)

R13591 HTLV-I env precursor epitope peptide.

ID R13591 standard; Protein; 40 AA.
AC R13591;
DT 03-OCT-1991 (first entry)
DE HTLV-I env precursor epitope peptide.
KW HTLV-I; epitope; diagnosis; env protein; gp46; antibody; vaccine.
OS Synthetic.
PN US5003043-A.
PD 26-MAR-1991.
PF 25-MAY-1988; 198416.
PE 17-MAY-1991; US-469291.

PR 24-JAN-1990; US-469291.
 PA (UNBI-) UNITED BIOMEDICAL.
 PI Wang CY;
 DR WPI: 91-224505/31.
 PT Peptide compsns. corresp. to envelope fragments of HTLV-1,2 - for
 PT detecting antibodies to these viruses and diagnosing HIV and
 PT adult T-cell leukaemia infections
 PS Claim 1; Page 17; 27pp; English.
 CC This peptide is one of 16 peptides useful for detecting antibodies to
 CC HTLV or HIV viruses. The peptides correspond to partial sequences of
 CC the HTLV virus designated gp21 and gp64, both part of gp61, which
 CC defines the envelope protein of the HTLV-I or HTLV-II virus. The
 CC peptides can be amidated at the C-terminal. This particular peptide
 CC is used in a composition with at least two of the other peptides of
 CC the invention. See R13184-R13193 and R13861-6.
 SQ Sequence 38 AA;
 SQ 2 A; 3 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 3 I; 4 L; 0 K; 0 M; 2 F; 6 P; 7 S; 1 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:31-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 GAEVS
 || ||
 CFDPQIGAIIVSSPCHNSLILPPFSLSPVPTLGSRSRRA
 X 10 20 30

10. US-08-249-182-4 (1-5)
 R06408 HTLV-1 corresponding peptide (VI).

ID R06408 standard; protein; 38 AA.
 AC R06408;
 DT 21-DEC-1990 (first entry)
 DE HTLV-1 corresponding peptide (VI).
 KW HTLV-1; HIV; antibodies; vaccines; polymers;
 OS Synthetic.
 PN W09008162-A.
 PD 26-JUL-1990.
 PF 16-JAN-1990; U00260.
 PR 13-JAN-1989; US-297635.
 PA (UNBI-) UNITED BIOMED INC.
 PI Yang CY;
 DR WPI: 90-254015/33.
 PT Synthetic peptide(s) corresponding to HTLV-1 and op. HIV - used
 PT for detection of antibodies, in vaccines and for development of
 PT antibodies
 PS Claim 1 (VI); Page 38; 52pp; English.
 CC Peptides having specific immunoreactivity to antibodies to HTLV-1
 CC comprise this sequence on its own, or an analogue of it in which
 CC amino acids may be added, deleted or substd, or segments, mixts.,
 CC conjugates or polymers of the peptides representes in R06403-08.
 CC The peptides are safe, sensitive and specific in the detection of
 CC antibodies. This peptide corresponds to a partial segment of the
 CC amino acid sequence of the HTLV-1 virus gp.21 or gp.46 and are
 CC prepared by solid phase synthesis.
 SQ Sequence 38 AA;
 SQ 2 A; 3 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 3 I; 4 L; 0 K; 0 M; 2 F; 6 P; 7 S; 1 T; 0 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:07-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74

ID diagnosis; prophylaxis; therapy; AIDS.
 OS Synthetic.
 PN W09003984-A.
 PD 19-APR-1990.
 PF 29-SEP-1989; U04302.
 PR 19-SEP-1989; US-407663, US-252949; W0-U04302.
 PA (REPL-) Repligen Corp.
 PI Rusche JR, Putney SD, Javaherian K, Farley J, Grimaldi R, Lynn D,
 PI Petro-Breyer J.
 DR WPI; 90-147824/19.
 PT Principal neutralising domain of HIV variants - used for producing
 PT peptide(s) and antibodies for diagnosis; prophylaxis; and/or therapy
 PT of HIV infection.
 PS Claim 8 (75); Page 77; 108pp; English.
 CC Peptide RP70 comprises segments of the Principal Neutralising Domain
 CC (envelope protein) from isolate HIV-MN. Cysteines may be added,
 CC so that the residues at or near both ends to form a disulfide bond,
 CC giving the peptide a loop-like configuration.
 CC The loop configuration can be utilised to enhance the
 CC immunogenic properties of the peptides. The protein is capable of
 CC eliciting, and/or binding with, neutralising antibodies.
 CC The neutralising domain is bounded by cysteine residues which occur at
 CC positions 296 and 331. Peptides can be used as immunogens or screening
 CC reagents to generate or identify poly- or monoclonal antibodies.
 CC See also R04427-R04506 and Q04273-Q04279.
 SQ Sequence 40 AA;
 SQ 2 A; 5 R; 5 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
 SQ 7 I; 0 L; 3 K; 0 M; 1 F; 2 P; 1 S; 4 T; 0 W; 2 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:56:55-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
    PEEVTRPNYL
      |||||
    INCTRPNYNKRRIHIGPGRFYTTKNIIGTIRGAHCNIS
    X      10      20      30      40
  
```

6. US-08-249-182-5 (1-10)

R31521 Cysteine protease #2.

ID R31521 standard; Protein; 215 AA.
 AC R31521;
 DT 20-MAY-1993 (first entry)
 DE Cysteine protease #2.
 KW Generic; cysteine protease; parasite; helminth; immune system;
 KW hypersensitivity reaction; antibody; antigen; graft; tissue;
 KW footpad; swelling; booster; sheep; red blood cell; immunisation;
 KW organ; transplant; rejection; autoimmune; disease.
 OS Paragonimus westermani.
 PN EP-524834-A.
 PD 27-JAN-1993.
 PF 24-JUL-1992; 306803.
 PR 25-JUL-1991; JP-208546.
 PR 12-FEB-1992; JP-057189.
 PA (HAMA/) HAMAJIMA F.
 PA (TSUR/) TSURU S.
 PA (YAMA/) YAMAKAMI K.
 PA (YAMA/) YAMAMOTO M.
 PI Hamajima F, Tsuru S, Yamakami K, Yanamoto M;
 DR WPI; 93-028881/04.
 DR N-PSDB; Q35446.
 PT Immunosuppressive cysteine protease obtained from larvae of

PT parasitic helminths - used for suppression of graft rejection in
 PS organ transplantation and suppression of auto-immune diseases
 PS Disclosure; Page 19-20; 26pp; English.
 CC The sequence given in R31519 is a generic version of the sequences
 CC given in R31520-21. These sequences represent a cysteine protease
 CC which was isolated from the parasitic helminth, *Paragonimus*
 CC *westermani*. This cysteine protease suppresses the immune system,
 CC both cell mediated and humoral. It inhibits delayed hypersensitivity
 CC reactions and production of antibodies against specific antigens and
 CC graft tissues. This protease inhibits footpad swelling caused by a
 CC booster injection of sheep red blood cells in the footpad after an
 CC initial immunisation intraperitoneally. The proteinase is useful for
 CC preventing organ transplant rejection and for control of autoimmune
 CC disease.
 SQ Sequence 215 AA;
 SQ 18 A; 5 R; 8 N; 14 D; 0 B; 8 C; 8 Q; 18 E; 0 Z; 21 G; 3 H;
 SQ 10 I; 17 L; 12 K; 5 M; 5 F; 9 P; 14 S; 11 T; 7 W; 9 Y; 13 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:38-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||    ||
GAWPSSSVLEIIDHGGLSESDYPYVGVEQTCALNKEKLVAKIDDSIVLGPEEEDHAAYLAEHGPLSTLLNA
  70      80      90      100      110  X 120  X 130

VALQYYQSGVLKPTFEPCDTELNHAULTVGYDKEGDM
  140      150      160      170

```

7. US-08-249-182-5 (1-10)

R31520 Cysteine protease #1.

ID R31520 standard; Protein; 215 AA.
 AC R31520;
 DT 20-MAY-1993 (first entry)
 DE Cysteine protease #1.
 KW Generic; cysteine protease; parasite; helminth; immune system;
 KW hypersensitivity reaction; antibody; antigen; graft; tissue;
 KW footpad; swelling; booster; sheep; red blood cell; immunisation;
 KW organ; transplant; rejection; autoimmune; disease.
 OS *Paragonimus westermani*.
 PN EP-524834-A.
 PD 27-JAN-1993.
 PF 24-JUL-1992; 306803.
 PR 25-JUL-1991; JP-208546.
 PR 12-FEB-1992; JP-057189.
 PA (YAMA/) HAMAJIMA F.
 PA (YAMA) TSURU S.
 PA (YAMA/) MAKAMI K.
 PI Hamajima F, TTD M.
 DR WPI; 93-028881/04 S, Yamakami K, Yamamoto M;
 DR N-PSDB; 035445.
 PT Immunosuppressive cysteine
 PT parasitic helminths - used for suppression of graft rejection in
 PT organ transplantation and suppression of auto-immune diseases
 PS Disclosure; Page 18; 26pp; English.
 CC The sequence given in R31519 is a generic version of the sequences
 CC given in R31520-21. These sequences represent a cysteine protease
 CC which was isolated from the parasitic helminth, *Paragonimus*
 CC *westermani*. This cysteine protease suppresses the immune system,
 CC both cell mediated and humoral. It inhibits delayed hypersensitivity

PEEVTRPNYL

||||

INCTRPNYNKRRIHIGPGRAFYTTKNIIGTIRGAHCNIS

X 10 20 30 40

4. US-08-249-182-5 (1-10)

R29228 Heteroconjugate antibody immunogen RP70.

ID R29228 standard; peptide; 40 AA.
AC R29228;
DT 14-APR-1993 (first entry)
DE Heteroconjugate antibody immunogen RP70.
KW V3 loop; gp41; envelope protein; MN; prototype; virus; variant; HIV;
KW homology; heteroconjugate; enzyme; epitope mapping; replication;
KW conjugate; immunogenic carrier; keyhole limpet hemocyanin; KLH;
KW ovalbumin; succinyl maleimidomethyl cyclohexanycarboxylate; SMCC.
OS Synthetic.
PN W09220373-A.
PD 26-NOV-1992.
PF 29-APR-1992; U03616.
PR 14-MAY-1991; US-699773.
PA (REPK) REPLIGEN CORP.
PI Higgins PJ, Potts BJ;
DR WPI; 92-415475/50.
PT Hetero-conjugate antibodies for treating HIV infections -
PT comprise an antibody specific for an effector cell surface
PT antigen and an antibody to a V3 loop of gp-120 envelope protein
PT of HIV
PS Disclosure; Page 19; 69pp; English.
CC The sequences given in R29226-35 represent peptides which were used
CC as immunogens for the production of antibodies against HIV. These
CC peptides may be either unconjugated or conjugated to an immunogenic
CC carrier, eg. a keyhole limpet hemocyanin (KLH) or ovalbumin, using
CC succinyl maleimidomethyl cyclohexanycarboxylate (SMCC) as a
CC conjugating agent. Viruses containing these or similar sequences may
CC be recognised by the heteroconjugate enzymes of the invention. The
CC antibodies raised against these sequences may be identified by standard
CC epitope mapping techniques. These antibodies are capable, even at low
CC concentrations, of nearly eliminating viral replication of different
CC strains of HIV.
SQ Sequence 40 AA;
SQ 2 A; 5 R; 5 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
SQ 7 I; 0 L; 3 K; 0 M; 1 F; 2 P; 1 S; 4 T; 0 W; 2 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:30-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
PEEVTRPNYL

||||

INCTRPNYNKRRIHIGPGRAFYTTKNIIGTIRGAHCNIS

X 10 20 30 40

5. US-08-249-182-5 (1-10)

R04461 Human immunodeficiency virus peptide RP70.

ID R04461 standard; protein; 40 AA.
AC R04461;
DT 20-SEP-1990 (first entry)
DE Human immunodeficiency virus peptide RP70.
KW HIV-MN; peptide RP70; principal neutralising domain; antibodies;

CC HIV-1 isolates having the amino acid sequence R33343. NM-01 is also
 CC putatively reactive with the RF-like peptide set out in R33344.
 CC The variable region of the heavy and light chain of monoclonal
 CC antibody NM-01 were cloned by PCR and sequenced. Nucleotides 1-21
 CC and 334-363 of Q37472 corresp. to the PCR primers used to amplify
 CC NM-01 light chain sequences and nucleotides 1-27 and 385-402 of
 CC Q57471 corresp. to the PCR primers used to amplify NM-01 heavy chain
 CC sequences.
 SQ Sequence 15 AA;
 SQ 0 A; 3 R; 2 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 2 I; 0 L; 2 K; 0 M; 0 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:45-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 71% Matches = 5 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

X      10
PEEVTRPNYL
|||||
CTRPNYNKRRIHIG
X      X 10

```

3. US-08-249-182-5 (1-10)

R33217 HIV gp120 V3 loop immunogenic peptide RP70.

ID R33217 standard; peptide; 40 AA.
 AC R33217;
 DT 13-JUL-1993 (first entry)
 DE HIV gp120 V3 loop immunogenic peptide RP70.
 KW HIV-1; human immunodeficiency virus; antibody generation; AIDS;
 KW infection; CD4 binding site; soluble CD4.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 16..20
 FT /note= "conserved subsequence in centre of V3 loop"
 PN W09304693-A.
 PD 18-MAR-1993.
 PF 02-SEP-1992; U07511.
 PR 09-SEP-1991; US-756677.
 PR 20-JUL-1992; US-916542.
 PA (REPK) REPLIGEN CORP.
 PI Field KG, Herlihy WC, Potts BJ, White-Scharf ME.
 DR WPI; 93-100653/12.
 PT Synergistic compsn. for treating HIV-1 infection - comprises antibody
 PT to V3 loop of gp120 and antibody to CD4 binding site of gp120 or
 PT soluble CD4 polypeptide
 PS Example; Page 12; 56pp; English.
 CC The sequence is that of peptide RP70 used as an immunogen for the
 CC generation of antibodies directed against the V3 loop of HIV gp120.
 CC These antibodies can be used as part of a compsn. with antibodies
 CC directed against the CD4 binding site of gp120. The antibodies act
 CC synergistically to neutralise HIV-1 in the treatment of HIV
 CC infection caused by different strains. It can be formed into a
 CC closed loop by creation of a disulphide bond between the two
 CC cysteine residues near the ends of the amino acid sequence.
 SQ Sequence 40 AA;
 SQ 2 A; 5 R; 5 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
 SQ 7 I; 0 L; 3 K; 0 M; 1 F; 2 P; 1 S; 4 T; 0 W; 2 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 29. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 10 AA;
 SQ 0 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 2 P; 0 S; 1 T; 0 W; 1 Y; 1 V;
 SQ 1 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 9 Optimized Score = 9 Significance = 6.66
 Residue Identity = 90% Matches = 9 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

X      X
PEEVTRPNYL
||||| |||
PEEVTXPNYL
X      10
  
```

2. US-08-249-182-5 (1-10)

R33332 Sequence of peptide which corresp. to residues 302

ID R33332 standard; peptide; 15 AA.
 AC R33332;
 DT 06-JUL-1993 (first entry)
 DE Sequence of peptide which corresp. to residues 302-316 of the V3
 DE loop region of HIV-1MN gp120.
 KW Monoclonal antibody; NM-01; HIV-1; gp120; gp160.
 OS Synthetic.
 PN W09304090-A.
 PD 04-MAR-1993.
 PF 24-AUG-1992; U07111.
 PR 22-AUG-1991; US-748562.
 PA (NISP) NISSIN SHOKUHI KAISHA LTD.
 PI Ohno T;
 DR WPI; 93-093943/11.
 PT Monoclonal antibodies against HIV-1 gp120 and gp160 proteins -
 PT for treating and preventing HIV-1 infection
 PS Example; Page 19; 57pp; English.
 CC Hybridoma cell line HB 10726 secretes MAb NM-01. In order to
 CC characterize the viral epitope recognized by NM-01, the Ab was
 CC screened reactivity with overlapping peptides corresp. to the amino
 CC acid sequence of the V3 loop region of HIV-1 gp120 (R33332, R33333,
 CC R33334). While there was no detectable reactivity over background of
 CC MAb-01 with the peptides corresp. to AAs 302-316 or 322-336 of the
 CC V3 loop, binding of the antibody to the peptide representing AAs
 CC 312-326 was apparent. The extent of this reactivity with other
 CC HIV-1 isolates was screened with peptides corresp. to the V3 loop
 CC region of HIV-1 isolates IIB, RF, CDC4, NY/5, Z6, Z2 and ELI
 CC (R33335-R33342). These results indicate that monoclonal antibody
 CC NM-01 recognizes an epitope of the V3 loop of gp120 of multiple

The list of best scores is:

		Init. Opt.			
Sequence Name	Description	Length	Score	Score	Sig. Frame

**** 6 standard deviations above mean ****					
1. R37447	Autotaxin peptide ATX 29.	10	9	9	6.66 0
**** 3 standard deviations above mean ****					
2. R33332	Sequence of peptide which cor	15	5	5	3.33 0
3. R33217	HIV gp120 V3 loop immunogenic	40	5	5	3.33 0
4. R29228	Heteroconjugate antibody immu	40	5	5	3.33 0
5. R04461	Human immunodeficiency virus	40	5	5	3.33 0
6. R31521	Cysteine protease #2.	215	5	5	3.33 0
7. R31520	Cysteine protease #1.	215	5	5	3.33 0
8. R31519	Cysteine protease generic seq	215	5	5	3.33 0
9. R44495	Sequence of the immunoglobuli	219	5	5	3.33 0
10. R04495	HIV fusion protein PB1mn	229	5	5	3.33 0
11. R07130	H20B receptor.	392	5	5	3.33 0
12. R07131	H20A receptor.	416	5	5	3.33 0
13. R14376	Factor XIII subunit a.	731	5	5	3.33 0
14. P70293	Sequence of human factor XIII	732	5	5	3.33 0
**** 2 standard deviations above mean ****					
15. R26038	Oligopeptide P24-34.	11	4	4	2.50 0
16. R36503	D32.39 antibody isolated pept	12	4	5	2.50 0
17. R04439	Human immunodeficiency virus	14	4	4	2.50 0
18. R04433	Human immunodeficiency virus	18	4	4	2.50 0
19. R24865	Sequence of peptide fragment	21	4	4	2.50 0
20. R30027	Gp120-29, for activating T ce	23	4	4	2.50 0
21. R22598	Beta tumour cell growth facto	24	4	4	2.50 0
22. R12179	HIV SP-10 region corresp. pep	24	4	4	2.50 0
23. R27325	Peptide corresp. to an epitop	25	4	4	2.50 0
24. P60885	Synthetic peptide which elici	25	4	4	2.50 0
25. R04435	Human immunodeficiency virus	25	4	4	2.50 0
26. R27324	Peptide corresp. to an epitop	26	4	4	2.50 0
27. R27318	Peptide corresp. to an epitop	26	4	4	2.50 0
28. P60887	Synthetic peptide which elici	26	4	4	2.50 0
29. R27319	Peptide corresp. to an epitop	27	4	4	2.50 0
30. R12180	HIV SP-10 region corresp. pep	30	4	4	2.50 0
31. R26867	HIV gp120SF2 V3 region derive	34	4	4	2.50 0
32. R26868	HIV gp120SF2 V3 region derive	36	4	4	2.50 0
33. R25747	gp120 epitope #2.	36	4	4	2.50 0
34. P83003	HIV envelope glycoprotein gp1	36	4	4	2.50 0
35. R41065	HIV-1 complete V3 loop sequen	37	4	4	2.50 0
36. R34286	HIV-1 isolate IIB env gp120	37	4	4	2.50 0
37. R34287	HIV-1 isolate IIB gp120 V3 l	38	4	4	2.50 0
38. R04428	Human immunodeficiency virus	39	4	4	2.50 0
39. R30020	HIV gp120 principle neutralis	41	4	4	2.50 0
40. R10749	Non-A non-B hepatitis specifi	85	4	4	2.50 0

1. US-08-249-182-5 (1-10)

R37447 Autotaxin peptide ATX 29.

ID R37447 standard; peptide; 10 AA.

AC R37447;

DT 22-JUL-1993 (first entry)

DE Autotaxin peptide ATX 29.

KW Cell motility stimulating; cancer metastasis; antibody; detection;

KW immunostains; disease outcome prediction; therapy choice;

KW cancer therapy; crosslinked toxins.

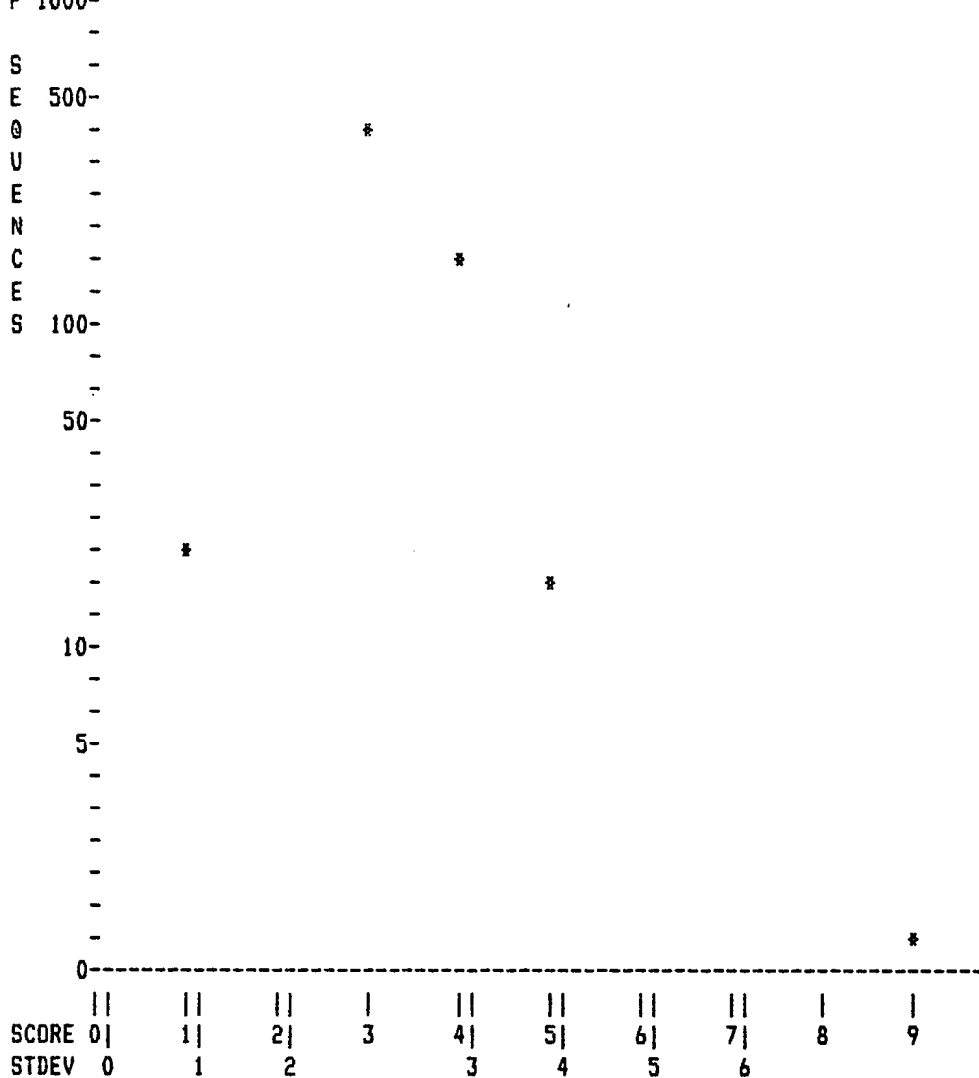
OS Synthetic.

FH Key Location/Qualifiers

FT Modified_site 6

FT /note= "potentially glycosylated residue"

PN US7822043-A.



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.20

Times:	CPU	Total Elapsed
	00:00:08.91	00:00:09.00

Number of residues:	482836
Number of sequences searched:	5543
Number of scores above cutoff:	2895

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

PR 23-OCT-1989; US-425252.
 PA (HOFF) Hoffmann-La Roche AG.
 PI Buonagurio DA, Longiaru M;
 DR WPI; 91-126308/18.
 DR N-PSDB; Q11554.
 PT New nucleic acid encoding conserved epitope of HTLV gp.21 - and
 PT hybrids with other epitopes and the derived polypeptide(s) useful as
 PT immunoassay reagents for detecting specific HTLV antibodies in serum
 PS Disclosure; fig 11; 31pp; English.
 CC This ENV93/HTLV-1-II fusion protein comprises an epitope from the
 CC immunodominant conserved region of HTLV-1 envelope (env) glyco-
 CC protein (gp)21 fused to an epitope from the gp46 region of HTLV-1
 CC env protein. The ENV93 epitope constitutes residues 342-432 of
 CC the gp21 sequence and the HTLV-1-II epitope constitutes residues
 CC 201-307 of gp46. The ENV93 gene construct encoding this fusion
 CC protein is is used as a vehicle for the high level expression
 CC of this other epitope of HTLV-1 env as a fusion protein. The
 CC fusion protein is useful for detecting antibodies to HTLV-1 in
 CC body fluids, eg blood, where it provides a more sensitive and
 CC selective assay than current viral lysate tests.
 CC See also Q11552-53 and Q11555-58.
 SQ Sequence 217 AA;
 SQ 10 A; 7 R; 11 N; 7 D; 0 B; 6 C; 14 Q; 6 E; 0 Z; 10 G; 7 H;
 SQ 11 I; 34 L; 8 K; 1 M; 5 F; 20 P; 25 S; 12 T; 6 W; 4 Y; 13 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:21-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.74
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                || ||
    ASLSTWHVLVSPNVSPSSSSTPLLYPSLALPAPHLTLFPNWTHCFDPQIGAIIVSSPCHNSLILPPFSLSPV
    140      150      160      170      180      190 X  200      210
  
```

PTLGSQA

> O <
 O| |O IntelliGenetics
 > O <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-08-249-182-5.res made by on Wed 21 Sep 94 12:09:17-PDT.

Query sequence being compared:US-08-249-182-5 (1-10)
 Number of sequences searched: 5543
 Number of scores above cutoff: 2895

Results of the initial comparison of US-08-249-182-5 (1-10) with:
 File : /home/shears/loring/lorin*.pep

10000-
 -
 N -
 U 5000-
 M -
 B *
 E - *
 R -
 -
 O -

X X
 YDVPWNETI
 ||||
 CTTAVPWNASWS
 X 10

7. US-08-249-182-6 (1-9)

P91670 HIV-1 TMP related polypeptide

ID P91670 standard; protein; 12 AA.
 AC P91670;
 DT 29-JUN-1990 (first entry)
 DE HIV-1 TMP related polypeptide
 KW HIV-1; TMP related polypeptide; assay; anti-HIV antibodies.
 OS Human immunodeficiency virus.
 PN W08901494-A.
 PD 23-FEB-1989.
 PF 19-AUG-1988; U02870.
 PR 21-AUG-1987; US-088067.
 PA (SCRI) Scripps Clinic & Res.
 PI Oldstone M, Gnann J;
 DR WPI; 89-068856/09.
 PT Human immuno-deficiency virus related polypeptide(s) -
 PT used for assaying anti-HIV antibodies in samples and for
 PT inducing antibody prodn.
 PS Disclosure; page 16; 43pp; English
 CC A compsn. is claimed comprising an admixture of at least two polypeptides
 CC where the 1st polypeptide is p91665 or p92120 and the 2nd is
 CC p91666, p91667 or p91668. The compsn. is characterised as being
 CC substantially free of antibodies (Abs) that immunoreact with HIV-1
 CC TMP related polypeptides such as p91670.
 SQ Sequence 12 AA;
 SQ 2 A; 0 R; 1 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:56:50-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 44% Matches = 4 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 ||||
 CTTAVPWNASWS
 X 10

8. US-08-249-182-6 (1-9)

R13191 Peptide (XIV) immunoreactive with anti-HIV-1 antib

ID R13191 standard; Protein; 19 AA.
 AC R13191;
 DT 09-OCT-1991 (first entry)
 DE Peptide (XIV) immunoreactive with anti-HIV-1 antibodies.
 KW human T-cell leukaemia virus; AIDS; ATL; detection;
 KW envelope protein gp61; acquired immunodeficiency syndrome;
 KW human immunodeficiency virus.
 OS Synthetic.
 PN EP-439077-A.
 PD 31-JUL-1991.
 PF 18-JAN-1991; 100616.
 PR 24-JAN-1990; US-469291.
 PA (UNBI-) UNITED BIOMEDICAL.

PT includes forming 1st and 2nd PCR admixtures, subjecting them to
 PT PCR thermo-cycles, sepg. double stranded DNA, hybridising, etc.
 PS Disclosure; Fig 6; 143pp; English.
 CC This sequence is the C terminal of R27568 encoded by a sequence
 CC inserted into lambda ZAP to give ImmunoZAP H expression vector. The VH
 CC or VL homologue is fused to the N terminal of this sequence.
 SQ Sequence 12 AA;
 SQ 0 A; 0 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 2 P; 2 S; 1 T; 0 W; 3 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:26-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 50% Matches = 4 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 ||||
 TSYPYDVPDYGS
 X 10 X

6. US-08-249-182-6 (1-9)

R23616 Peptide able to induce in vivo prodn. of antibody

ID R23616 standard; peptide; 12 AA.
 AC R23616;
 DT 29-OCT-1992 (first entry)
 DE Peptide able to induce in vivo prodn. of antibodies.
 KW Mycoplasma; adhesion; molecule; HIV; AIDS; gp160; nef; T-lymphocytes;
 KW vaccine; passive; immunotherapy.
 OS Human immunodeficiency virus.
 PN W09206199-A.
 PD 16-APR-1992.
 PF 27-SEP-1991; F00758.
 PR 27-SEP-1990; FR-011951.
 PR 26-OCT-1990; FR-013324.
 PA (INSP) INST PASTEUR.
 PI Bahrooui E, Berneman D, Blanchard A, Chamaret S, Guetard D;
 PI Montagnier L, Van Rietschoten J;
 DR WPI; 92-150884/18.
 PT New peptide(s) common to mycoplasma adhesion molecule and HIV
 PT proteins - useful as vaccines against AIDS and mycoplasma
 PT infections and in their diagnosis
 PS Claim 21; Page 36; 123pp; French.
 CC The peptide is derived from the HIV envelope protein gp160 and has
 CC homology to a peptide derived from mycoplasma genitalium, a
 CC mycoplasma adhesin. The synthetic peptide, coupled to keyhole
 CC limpet haemocyanin to form a conjugate, is useful as an
 CC immunogen to induce antibodies which neutralise HIV infection.
 CC It can be formulated in a vaccine, opt. with other peptides able
 CC to induce neutralising antibodies (specifically one corresp. to
 CC the V3 region of gp120). The peptide can also be used diagnostic-
 CC ally to detect antibodies against mycoplasma and/or HIV and also
 CC to monitor the progress of disease treatment. Antibodies raised
 CC against the peptide may be used in passive immunotherapy to inhibit
 CC infection by HIV-T cell lymphocytes.
 CC See also R23604-17.
 SQ Sequence 12 AA;
 SQ 2 A; 0 R; 1 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 0 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:04-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 44% Matches = 4 Mismatches = 5

FT /note= "N-linked glycosylation site" 306..308
 FT Modified_site 306..308
 FT /note= "N-linked glycosylation site" 356..358
 FT Modified_site 356..358
 FT /note= "N-linked glycosylation site" 401..403
 FT Modified_site 401..403
 FT /note= "N-linked glycosylation site" 433..453
 FT Domain 433..453
 FT /note= "Transmembrane domain"
 PN USS237051-A.
 PD 17-AUG-1993.
 PR 06-DEC-1990; 623033.
 PR 06-DEC-1990; US-623033.
 PA (UYVA-) UNIV VANDERBILT.
 PI Garbers DL, Schulz S;
 DR MPI; 93-272183/34.
 DR P-PSDB; R38861.
 PT New purified enterotoxin receptor protein - used to develop
 PT prods. for treating abnormal conditions caused by bacterially
 PT released enterotoxin, partic. diarrhoea
 PS Claim 2; Fig 1; 26pp; English.
 CC This sequence represents guanylyl cyclase, GC-C, which binds heat
 CC stable enterotoxin. The DNA encoding this protein was isolated from
 CC rat small intestinal mucosa polyA+ RNA by PCR. This protein is an
 CC enterotoxin receptor which may be used as a therapeutic to control
 CC intestinal fluid permeation as well as abnormal conditions caused
 CC by bacterially released enterotoxin. The binding domain of the
 CC protein, or antibodies to the protein, can be used to eliminate
 CC diarrhoea. The protein may be used to isolate ligands and to screen
 CC for antagonists of toxin binding.
 SS Sequence 1075 AA;
 SS 42 A; 64 R; 48 N; 68 D; 0 B; 20 C; 32 Q; 73 E; 0 Z; 55 G; 23 H;
 SS 57 I; 118 L; 68 K; 32 M; 52 F; 43 P; 75 S; 71 T; 11 W; 48 Y; 75 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.59
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 || ||
 || ||
 TTPKFARAFRNLTFGLEGCPVTLDDSGDIDNIMHCLLVSLDTRKVKVLMAYDTHKNQTIIPVATSPNFIWMKNH
 350 360 370 380 390 X 400 X 410
 RLPNDVPGLGPPQILMIAVFTLTGIVVLLLIALLLR
 420 430 440 450

5. US-08-249-182-6 (1-9)
 R27567 C terminal end of sequence inserted into lambda ZA

ID R27567 standard; Protein: 12 AA.
 AC R27567;

DT 26-FEB-1993 (first entry)
 DE C terminal end of sequence inserted into lambda ZAP giving ImmunozAP
 KM Dictionic expression vector; fusion PCR; antibody; cDNA library;
 KM ss.
 PN M09215678-A.
 PD 17-SEP-1992.
 PF 27-FEB-1992; U01475.
 PR 01-MAR-1991; US-663442.
 PA (STRA-) STRATAGENE.
 PI Sorge JA;
 DR MPI; 92-331724/40.
 DR P-PSDB; 029216.

KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
 OS Rattus rattus.
 PN US5237051-A.
 PD 17-AUG-1993.
 PF 06-DEC-1990; 623033.
 PR 06-DEC-1990; US-623033.
 PA (UYVA-) UNIV VANDERBILT.
 PI Garbers DL, Schulz S;
 DR WPI; 93-272183/34.
 PT New purified enterotoxin receptor protein - used to develop
 PT prods. for treating abnormal conditions caused by bacterially
 PT released enterotoxin, partic. diarrhoea
 PS Disclosure; Fig 3; 26pp; English.
 CC The sequences given in R38862-63 represent the guanylyl cyclases,
 CC GC-A and GC-B, which binds heat stable enterotoxin. These proteins
 CC are enterotoxin receptors which may be used as a therapeutic to control
 CC intestinal fluid permeation as well as abnormal conditions caused
 CC by bacterially released enterotoxin. The binding domain of the
 CC proteins, or antibodies to the proteins, can be used to eliminate
 CC diarrhoea. The proteins may be used to isolate ligands and to screen
 CC for antagonists of toxin binding.
 SQ Sequence 1029 AA;
 SQ 74 A; 73 R; 37 N; 56 D; 0 B; 16 C; 40 Q; 72 E; 0 Z; 75 G; 23 H;
 SQ 38 I; 124 L; 44 K; 20 M; 46 F; 50 P; 61 S; 53 T; 16 W; 34 Y; 77 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.59
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||  |||
ETVQAEAFDSVTIYFSDIVGFTALSAESTPMQVVTLNDLYTCFDAVIDNFDVYKVETIGDAYMVVSGLPVR
      840      850      860      870      880 X      890      900

NGQLHAREVARMALALLDAVRSFRIRHRPQEQRLRLRI
      910      920      930      940
  
```

4. US-08-249-182-6 (1-9)
 R38861 GC-C.

ID R38861 standard; Protein; 1075 AA.
 AC R38861;
 DT 08-FEB-1994 (first entry)
 DE GC-C.
 KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
 KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
 KW binding domain; antibody; diarrhoea; ligand; antagonist.
 OS Rattus rattus.
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /note= "Signal peptide"
 FT Protein 23..1075
 FT /note= "Mature GC-C"
 FT Modified_site 31..33
 FT /note= "N-linked glycosylation site"
 FT Modified_site 74..76
 FT /note= "N-linked glycosylation site"
 FT Modified_site 78..80
 FT /note= "N-linked glycosylation site"
 FT Modified_site 187..189
 FT /note= "N-linked glycosylation site"
 FT Modified_site 194..196

```

X      X
YDVPWNETI
|||||||
YDVPWNETI
X      X

```

2. US-08-249-182-6 (1-9)
R38863 GC-B.

ID R38863 standard; Protein; 1025 AA.
AC R38863;
DT 08-FEB-1994 (first entry)
DE GC-B.
KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;
KW mucosa; polyA+ RNA; PCR; enterotoxin receptor; bacterial enterotoxin;
KW binding domain; antibody; diarrhoea; ligand; antagonist.
OS Rattus rattus.
PN US5237051-A.
PD 17-AUG-1993.
PF 06-DEC-1990; 623033.
PR 06-DEC-1990; US-623033.
PA (UYVA-) UNIV VANDERBILT.
PI Garbers DL, Schulz S;
DR WPI; 93-272183/34.
PT New purified enterotoxin receptor protein - used to develop
PT prods. for treating abnormal conditions caused by bacterially
PT released enterotoxin, partic. diarrhoea
PS Disclosure; Fig 3; 26pp; English.
CC The sequences given in R38862-63 represent the guanylyl cyclases,
CC GC-A and GC-B, which binds heat stable enterotoxin. These proteins
CC are enterotoxin receptors which may be used as a therapeutic to control
CC intestinal fluid permeation as well as abnormal conditions caused
CC by bacterially released enterotoxin. The binding domain of the
CC proteins, or antibodies to the proteins, can be used to eliminate
CC diarrhoea. The proteins may be used to isolate ligands and to screen
CC for antagonists of toxin binding. This sequence is given as it is
CC represented in the specification.
SQ Sequence 1025 AA;
SQ 82 A; 69 R; 39 N; 54 D; 0 B; 15 C; 40 Q; 64 E; 0 Z; 78 G; 28 H;
SQ 49 I; 122 L; 39 K; 21 M; 46 F; 53 P; 56 S; 53 T; 16 W; 37 Y; 64 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 4.59
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||  |||
ETVQAEAFDSVTIYFSDIVGFTALSAESTPMQVVTLNDLYTCFDAIIDNFVYKVGETIGDAYMVVSGLPGR
830      840      850      860      870      880      X 890

NGQRHAFEIARMALALLDAVSSFRIRHRPHDQLRLRI
900      910      920      930

```

3. US-08-249-182-6 (1-9)
R38862 GC-A.

ID R38862 standard; Protein; 1029 AA.
AC R38862;
DT 08-FEB-1994 (first entry)
DE GC-A.
KW Guanylyl cyclase; GC-C; heat stable; enterotoxin; rat; small intestine;

14. R02714	Peptide for detection of anti	34	4	4	3.67	0
15. R44497	Sequence of the HIV-1 epitope	35	4	4	3.67	0
16. R31964	BCH-87ck.	35	4	4	3.67	0
17. R31963	BCH-87c.	35	4	4	3.67	0
18. R26492	Fragment of HIV-1 gp41 (resid	35	4	4	3.67	0
19. R31966	BCH-266.	36	4	4	3.67	0
20. R31967	BCH-408.	39	4	4	3.67	0
21. R20661	HIV antibody detecting peptid	39	4	4	3.67	0
22. R31968	BCH-408k.	40	4	4	3.67	0
23. R20659	HIV antibody detecting peptid	44	4	4	3.67	0
24. P60067	Sequence of HTLV-III polypept	82	4	5	3.67	0
25. P70239	Polypeptide ENV(80).	96	4	4	3.67	0
26. R25407	Light chain variable domain o	113	4	4	3.67	0
27. R28743	Light chain variable domain o	114	4	4	3.67	0
28. P93342	HIV-1 env protein from CBL-1	140	4	4	3.67	0
29. P60068	Sequence of HTLV-III polypept	146	4	4	3.67	0
30. R41320	PEP (1-160).	160	4	4	3.67	0
31. R43910	Nerve growth factor.	239	4	4	3.67	0
32. R37647	Sequence of a 4-4-20/212 sing	250	4	4	3.67	0
33. R30723	Protein C heavy chain.	261	4	4	3.67	0
34. R34030	Fc-alpha-R.	287	4	4	3.67	0
35. P80297	Sequence encoded by the env (291	4	4	3.67	0
36. P82053	Outer membrane protein F of P	350	4	4	3.67	0
37. R05637	Placenta-specific protein PP1	369	4	4	3.67	0
38. R21846	Gal alpha-2,6-ST (from clone	406	4	4	3.67	0
39. P91235	(ENV-80)(GAG-VII)(Hexahis) pr	515	4	4	3.67	0
40. P70541	HTLV-III gag/env gene protein	600	4	4	3.67	0

1. US-08-249-182-6 (1-9)

R37448 Autotaxin peptide ATX 47.

ID R37448 standard; peptide; 9 AA.
AC R37448;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 47.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 47. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 9 AA;
SQ 0 A; 0 R; 1 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 1 I; 0 L; 0 K; 0 M; 0 F; 1 P; 0 S; 1 T; 1 W; 1 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 9 Optimized Score = 9 Significance = 8.27
Residue Identity = 100% Matches = 9 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

*

*

```

0-----
||  ||  ||  ||  ||  ||  |  ||  |  |
SCORE 0| 1| 2| 3| 4| 5| 6 7| 8 9
STDEV 1  2  3  4  5  6  7

```

PARAMETERS

```

Similarity matrix      Unitary      K-tuple      2
Mismatch penalty      1      Joining penalty      20
Gap penalty            1.00      Window size      5
Gap size penalty      0.05
Cutoff score          0
Randomization group   0

```

```

Initial scores to save      40      Alignments to save      15
Optimized scores to save    0      Display context      50

```

SEARCH STATISTICS

```

Scores:                Mean      Median      Standard Deviation
                        0          1          1.09

```

```

Times:                CPU          Total Elapsed
                   00:00:08.90      00:00:09.00

```

```

Number of residues:      482836
Number of sequences searched: 5543
Number of scores above cutoff: 2303

```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37448	Autotaxin peptide ATX 47.	9	9	9	8.27	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
2. R38863	GC-B.	1025	5	5	4.59	0
3. R38862	GC-A.	1029	5	5	4.59	0
4. R38861	GC-C.	1075	5	5	4.59	0
**** 3 standard deviations above mean ****						
5. R27567	C terminal end of sequence in	12	4	4	3.67	0
6. R23616	Peptide able to induce in viv	12	4	4	3.67	0
7. P91670	HIV-1 TMP related polypeptide	12	4	4	3.67	0
8. R13191	Peptide (XIV) immunoreactive	19	4	4	3.67	0
9. R06410	HTLV-1 corresponding peptide	19	4	4	3.67	0
10. R41059	HIV-1 gp41 peptide (isolate H	21	4	4	3.67	0
11. R41063	HIV-1 gp41 peptide (isolate E	25	4	4	3.67	0
12. P80566	Peptide region of human immun	27	4	4	3.67	0
13. R27564	Insert D to prevent steric hi	29	4	4	3.67	0

CC Dependent neoplastic disease, eg. lymphoma or leukaemia.
 SQ Sequence 11 AA;
 SQ 0 A; 1 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 1 H;
 SQ 2 I; 0 L; 1 K; 0 M; 0 F; 1 P; 0 S; 1 T; 1 W; 0 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:21-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.50
 Residue Identity = 44% Matches = 4 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

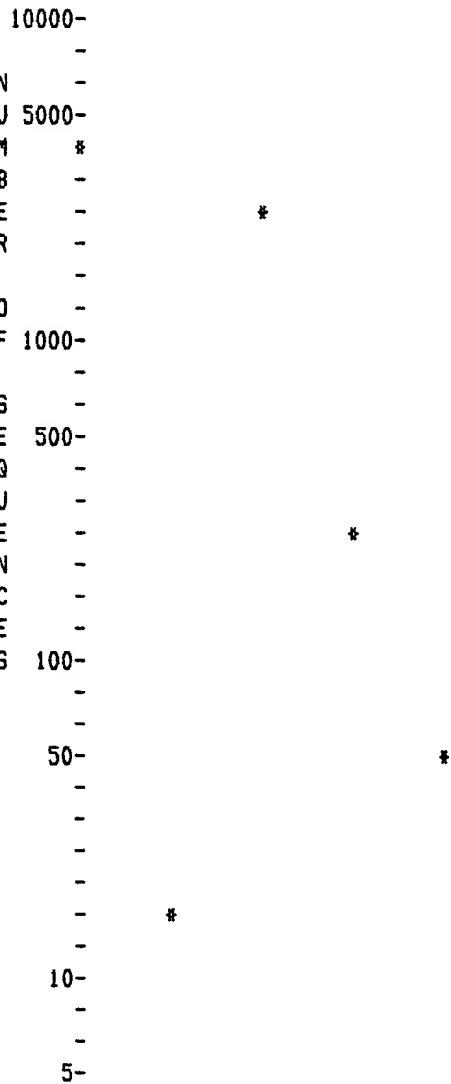
X X
 PEEVTRPNYL
 ||||
 IIVTRPWKHVE
 X 10
 > 0 <
 0| 0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-08-249-182-6.res made by on Wed 21 Sep 94 12:10:24-PDT.

Query sequence being compared:US-08-249-182-6 (1-9)
 Number of sequences searched: 5543
 Number of scores above cutoff: 2303

Results of the initial comparison of US-08-249-182-6 (1-9) with:
 File : /home/shears/loring/lorin*.pep



PR 12-MAR-1986; DE-608280.
 PR 26-JUN-1986; DE-621371.
 PA (BEHW) BEHRINGWERKE AG.
 PI Grundmann U, Amann E, Zettlmeissl G;
 DR WPI; 87-258275/37.
 DR N-PSDB; N70461.
 PT New DNA sequence coding for factor 13A and expressed proteins -
 PT useful as diagnostic reagents and for producing antibodies
 PS Claim 10; Table 3, ppl6-20; 30pp; German.
 CC Human placental cDNA gene bank was screened by hybridisation with two
 CC synthetic oligonucleotides, corresponding to partial AA sequences of
 CC factor XIIIa (N70460, N70465). N70461 gives the coding strand
 CC sequence of clones lambda-gt10-11 and lambda-gt10-12.
 SQ Sequence 732 AA;
 SQ 37 A; 45 R; 40 N; 47 D; 0 B; 9 C; 26 Q; 49 E; 0 Z; 50 G; 14 H;
 SQ 38 I; 49 L; 38 K; 20 M; 31 F; 33 P; 46 S; 45 T; 15 W; 29 Y; 71 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:14-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                PEEVTRPNYL
                                ||| ||
VVGSDMTVTVQFTNPLKETLRNVVHLDGPGVTRPMKKMFREIRPNSTVQWEEVCRPWVSGHRKLIASMSSD
      650      660      670      680      690 X      700      710

SLRHVYGELDVQIGRRPSM
      720      730
  
```

15. US-08-249-182-5 (1-10)
 R26038 Oligopeptide P24-34.

ID R26038 standard; peptide; 11 AA.
 AC R26038;
 DT 02-FEB-1993 (first entry)
 DE Oligopeptide P24-34.
 KW Human; granulocyte-macrophage colony stimulating factor; antibody;
 KW monoclonal; haemopoietic progenitor cells; neoplastic disease;
 KW lymphoma; leukaemia.
 OS Mus musculus.
 PN EP-499161-A.
 PD 19-AUG-1992.
 PF 07-FEB-1992; 102103.
 PR 11-FEB-1991; US-653428.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Braslawsky GR, Bursuker I, Greenfield RS;
 DR WPI; 92-277909/34.
 PT New monoclonal antibodies which inhibit GM-CSF - for treating
 PT GM-CSF mediated inflammatory and auto:immune diseases e.g.
 PT systemic lupus erythematosus, temporal arteritis, etc., also
 PT atherosclerosis
 PS Disclosure; Fig 4; 31pp; English.
 CC The sequence given in R26036 is a portion of the murine wild-type
 CC sequence from the granulocyte-macrophage colony stimulating factor
 CC (GM-CSF). The sequence given in R26037 is the "corresponding"
 CC sequence from the human GM-CSF polypeptide sequence. Sequences
 CC R26038-41 are other portions from the murine GM-CSF gene. All these
 CC peptide sequences can be used to raise antibodies against GM-CSF.
 CC These monoclonal antibodies (MAb) can be used to inhibit or prevent
 CC undesirable disease states which are caused by the colony stimulating
 CC activity of GM-CSF on haemopoietic progenitor cells. These MAb's can
 CC also be used for inhibiting or eliminating the growth of GM-CSF-

FT Protein 38..731
 FT /label= mature_a' subunit
 FT Active_site 314
 FT /note= "of the a subunit"
 FT Cleavage_site 37..38
 FT /note= "for conversion of a subunit to a' subunit
 FT by thrombin"
 PN W09116931-A.
 PD 14-NOV-1991.
 PF 09-MAY-1991; U03212.
 PR 10-MAY-1990; US-521805.
 PR 18-MAY-1990; US-525556.
 PA (ZYMO-) ZYMOGENETICS INC.
 PI Bishop PD;
 DR WPI; 91-353537/48.
 DR N-PSDB; Q14687.
 PT New diagnostic compsns. contg. factor XIII or derivs. -
 PT are coupled to radioisotope or paramagnetic cpd. and are useful
 PT for detecting thrombosis in high risk patients
 PS Disclosure; Fig 1; 51pp; English.
 CC A diagnostic compsn. comprising factor XIII couples to a detectable
 CC substance, and a physiologically acceptable carrier, are used to
 CC detect and image blood clots (venous and arterial thrombosis).
 CC Since they have a longer half life than fibrinogen, they are
 CC partic. useful before surgery to provide on-going monitoring (for up
 CC to 2 weeks after surgery) from a single dose.
 CC Mutant forms of factor XIII were made by in vitro site-directed
 CC mutagenesis, esp. ss pRS202 templates were modified by the single
 CC primer method, e.g. for TGC-Cys-314 to TCT-Ser exchange.
 CC See also Q14687-88.
 SQ Sequence 731 AA;
 SQ 37 A; 45 R; 40 N; 47 D; 0 B; 9 C; 27 Q; 48 E; 0 Z; 50 G; 14 H;
 SQ 39 I; 49 L; 38 K; 19 M; 31 F; 33 P; 46 S; 45 T; 15 W; 29 Y; 70 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:40-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 PEEVTRPNYL
 ||| ||
 VVGSDMTVTIQFTNPLKETLRNVVHLDGPGVTRPMKKMFREIRPNSTVQWEEVCRPWVSGHRKLIASMSSD
 650 660 670 680 690 700 710
 SLRHVYGELDVQIQRRPSM
 720 730

14. US-08-249-182-5 (1-10)

P70293 Sequence of human factor XIIIa.

ID P70293 standard; Protein; 732 AA.
 AC P70293;
 DT 21-MAY-1991 (first entry)
 DE Sequence of human factor XIIIa.
 KW Diagnosis; antigen; anti-factor XIIIa antibody.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 228..249
 FT /note= "used to derive 66mer probe N70460"
 FT Region 475..481
 FT /note= "used to derive 20mer probe N70465"
 PN EP-236978-A.
 PD 16-SEP-1987.
 PF 09-MAY-1991; U03212.

12. US-08-249-182-5 (1-10)

R07131 H20A receptor.

ID R07131 standard; protein; 416 AA.
 AC R07131;
 DT 23-JAN-1991 (first entry)
 DE H20A receptor.
 KW Picornavirus proteins; poliovirus; transgenic animals; vaccines;
 KW antibodies; imaging.
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label=signal peptide
 FT Domain 345..368
 FT /label=transmembrane domain
 PN W09010699-A.
 PD 20-SEP-1990.
 PF 09-MAR-1990; U01320.
 PR 10-MAR-1989; US-321957.
 PA (UYCO-) COLUMBIA UNIV NY.
 PI Racaniello V, Mendelsohn C, Costantini F;
 DR WPI; 90-305023/40.
 DR N-PSDB; 006070.
 PT DNA encoding picornavirus partic. poliovirus receptors proteins -
 PT for treating picornavirus infections or for expression in
 PT transgenic animals used to test vaccines
 PS Disclosure; fig 4; 88pp; English.
 CC This poliovirus receptor, H20A, has a sequence differing from the
 CC receptor, H20B at the cytoplasmic tail only. Antibodies (Abs)
 CC raised against it are useful for targetted delivery of the human
 CC poliovirus, conjugated to a drug. Transgenic animals contg. the
 CC corresp. DNA (genomic- or cDNA) can be used to test the efficiency
 CC and virulence of picornavirus vaccines.
 CC See also 006069.
 SQ Sequence 416 AA;
 SQ 30 A; 18 R; 18 N; 10 D; 0 B; 9 C; 21 Q; 22 E; 0 Z; 32 G; 10 H;
 SQ 13 I; 42 L; 9 K; 9 M; 12 F; 32 P; 37 S; 31 T; 10 W; 11 Y; 40 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:08-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                X      X
              PEEVTRPNYL
                || || ||
MARAMAAAWPLLLVALLVLSWPPPGTGDVVVQAPTQVPGFLGDSVTLPCYLQVPNMEVTHVSQLTWARHGES
      10      20      30      40 X      50      60      70

GSMVAFHQQTQGPSYSESKRLEFVAARLGA
      80      90     100

```

13. US-08-249-182-5 (1-10)

R14376 Factor XIII subunit a.

ID R14376 standard; Protein; 731 AA.
 AC R14376;
 DT 14-FEB-1992 (first entry)
 DE Factor XIII subunit a.
 KW Factor XIII; subunit; antibody; diagnosis; thrombin; thrombosis;
 KW mutant.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Protein 1..731
 FT

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
PVETPTREIKKLDGLWAFSLDRERVADIIDGSENFDTNAKTIIVHLNESVQINCTRPNYNKRKRIHIGPGRA
  10      20      30      40      50 X 60 X 70

FYTTKNIIGTIRGAHCNISRAKWNQTLRGIVSKLKEQF
  80      90      100     110
  
```

11. US-08-249-182-5 (1-10)

R07130 H20B receptor.

ID R07130 standard; protein; 392 AA.
 AC R07130;
 DT 23-JAN-1991 (first entry)
 DE H20B receptor.
 KW Picornavirus proteins; poliovirus; transgenic animals; vaccines;
 KW antibodies; imaging.
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label=signal peptide
 FT Domain 345..368
 FT /label=transmembrane domain
 PN W09010699-A.
 PD 20-SEP-1990.
 PF 09-MAR-1990; U01320.
 PR 10-MAR-1989; US-321957.
 PA (UYCO-) COLUMBIA UNIV NY.
 PI Racaniello V, Mendelsohn C, Costantini F;
 DR WPI; 90-305023/40.
 DR N-PSDB; 006069.
 PT DNA encoding picornavirus partic. poliovirus receptors proteins -
 PT for treating picornavirus infections or for expression in
 PT transgenic animals used to test vaccines
 PS Disclosure; fig 4; 88pp; English.
 CC This poliovirus receptor, H20B, has a sequence differing from the
 CC receptor, H20A at the cytoplasmic tail only. Antibodies (Abs)
 CC raised against it are useful for targetted delivery of the human
 CC poliovirus, conjugated to a drug. Transgenic animals contg. the
 CC corresp. DNA (genomic- or cDNA) can be used to test the efficiency
 CC and virulence of picornavirus vaccines.
 CC See also 006070.
 SQ Sequence 392 AA;
 SQ 26 A; 17 R; 17 N; 9 D; 0 B; 10 C; 20 Q; 20 E; 0 Z; 30 G; 10 H;
 SQ 13 I; 42 L; 9 K; 9 M; 12 F; 31 P; 31 S; 28 T; 10 W; 10 Y; 38 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:08-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                PEEVTRPNYL
                                |||||
MARAAAAWPLLLVALLVLSWPPPGTGDVVVQAPTQVPGFLGDSVTLPCYLQVPNMEVTHVSQLTWARHGES
  10      20      30      40 X 50      60      70

GSMVVFHGTQGPSYSESKRLEFVAARLGA
  80      90      100
  
```

PS Claim 6; Page 13-14; 26pp; English.
 CC The sequence given in R31519 is a generic version of the sequences
 CC given in R31520-21. These sequences represent a cysteine protease
 CC which was isolated from the parasitic helminth, *Paragonimus*
 CC *westerni*. This cysteine protease suppresses the immune system,
 CC both cell mediated and humoral. It inhibits delayed hypersensitivity
 CC reactions and production of antibodies against specific antigens and
 CC graft tissues. This protease inhibits footpad swelling caused by a
 CC booster injection of sheep red blood cells in the footpad after an
 CC initial immunisation intraperitoneally. The proteinase is useful for
 CC preventing organ transplant rejection and for control of autoimmune
 CC disease.
 SQ Sequence 215 AA;
 SQ 17 A; 5 R; 8 N; 12 D; 0 B; 8 C; 8 Q; 17 E; 0 Z; 21 G; 3 H;
 SQ 10 I; 17 L; 12 K; 4 M; 5 F; 8 P; 13 S; 11 T; 7 W; 9 Y; 13 V;
 SQ 7 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:38-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||    ||
GGWPXSSYLEIMXMGGLESEDYPYVGVE@TCALNKEKLVAKIDDSIVLGP EEEDHAAYLAEHGPLSTLLNA
   70      80      90      100      110  X 120  X 130

VALQYYQSGVLKPTFE ECPDTELNHA VLT VGYDKEGDM
   140     150     160     170
  
```

9. US-08-249-182-5 (1-10)

R44495 Sequence of the immunoglobulin IgG (1C3/86) kappa

ID R44495 standard; Protein; 219 AA.
 AC R44495;
 DT 26-MAY-1994 (first entry)
 DE Sequence of the immunoglobulin IgG (1C3/86) kappa chain
 DE derived from clone gamma1.1.1a and augmented using PCR.
 KW Particle-binding antibody fragment; kappa chain;
 KW monoclonal cell line 1C3/86; anti-erythrocyte IgGs.
 OS Synthetic.
 PN W09324630-A.
 PD 09-DEC-1993.
 PF 19-MAY-1993; AU0228.
 PR 22-MAY-1992; AU-002551.
 PA (AGEN-) AGEN LTD.
 PI Hillyard CJ, Hudson PJ, Lilley GG;
 DR WPI; 93-405821/50.
 DR N-PSDB; Q53430.
 PT Bifunctional recombinant protein - contains particle and analyte
 PT binding moieties, used in agglutination assays pref. on whole
 PT blood
 PS Example; Figure 2; 42pp; English.
 CC mRNA was prep'd. from monoclonal antibody cell line (1C3/86) which
 CC prods anti-erythrocyte IgGs which bind to RBCs. ds-cDNA was prep'd.
 CC and cloned into lambda-gt10 arms and packaged into a phage library.
 CC The heavy chain clone gamma-M/1.1 and the light chain clone
 CC ph76-kappa-10 were used to source ds-DNA inserts for the screening
 CC of the gt10 library. Positive clones were amplified, and the positive
 CC insert cDNA subcloned into pUC18. As a result a near full-length gamma
 CC clone (gamma-1.1.1a) was identified, the nucleotide sequence was
 CC determined and from this the protein sequence deduced (Q53429/R44494).
 CC The sequences of a partial kappa clone (kappa-4AC1) which encoded

CC graft tissues. This protease inhibits footpad swelling caused by a
 CC booster injection of sheep red blood cells in the footpad after an
 CC initial immunisation intraperitoneally. The proteinase is useful for
 CC preventing organ transplant rejection and for control of autoimmune
 CC disease.
 SQ Sequence 215 AA;
 SQ 18 A; 5 R; 8 N; 12 D; 0 B; 7 C; 9 Q; 17 E; 0 Z; 23 G; 3 H;
 SQ 10 I; 17 L; 12 K; 4 M; 5 F; 8 P; 15 S; 11 T; 7 W; 10 Y; 14 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:38-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||  ||
GGWPASSYLEIMYNGGLESESDYPYVGVEQTCALNKEKLVAKIDDSIVLGPEEEDHAAYLAEHGPLSTLLNA
  70      80      90      100      110  X 120  X 130

VALQYYQSGVLKPTFEESPDTELNHAULTVGYDKEGDM
  140      150      160      170

```

8. US-08-249-182-5 (1-10)

R31519 Cysteine protease generic sequence.

ID R31519 standard; Protein; 215 AA.
 AC R31519;
 DT 20-MAY-1993 (first entry)
 DE Cysteine protease generic sequence.
 KW Generic; cysteine protease; parasite; helminth; immune system;
 KW hypersensitivity reaction; antibody; antigen; graft; tissue;
 KW footpad; swelling; booster; sheep; red blood cell; immunisation;
 KW organ; transplant; rejection; autoimmune; disease.
 OS Paragonimus westermani.
 FH Key Location/Qualifiers
 FT Misc_difference 15
 FT /label= Ala, Pro
 FT Misc_difference 21
 FT /label= Ser, Glu
 FT Misc_difference 58
 FT /label= Arg, Met
 FT Misc_difference 59
 FT /label= Val, Ala
 FT Misc_difference 61
 FT /label= Gln, Glu
 FT Misc_difference 69
 FT /label= Ala, Ser
 FT Misc_difference 77
 FT /label= Tyr, Asp
 PN EP-524834-A.
 PD 27-JAN-1993.
 PF 24-JUL-1992; 306803.
 PR 25-JUL-1991; JP-208546.
 PR 12-FEB-1992; JP-057189.
 PA (HAMA/) HAMAJIMA F.
 PA (TSUR/) TSURU S.
 PA (YAMA/) YAMAKAMI K.
 PA (YAMA/) YAMAMOTO M.
 PI Hamajima F, Tsuru S, Yamakami K, Yamamoto M;
 DR WPI; 93-028881/04.
 DR N-PSDB; Q35444.
 PT Immunosuppressive cysteine protease obtained from larvae of
 PT parasitic helminths - used for suppression of graft rejection in

CC determined in a similar fashion. To determine the sequence of the
 CC 1C3/86 kappa light chain at the 5' end, a mixed N-terminal sequence
 CC was determined from the intact 1C3/86 Ig and together with the
 CC sequence from a gamma heavy chain clone used to determine the
 CC N-terminus of the variable region of the kappa light chain. A
 CC coding sequence for this amino acid sequence was compiled and PCR
 CC amplified using the redundant forward (sense) primer N960 and the
 CC reverse (antisense) primer N852) which was based on the kappa
 CC constant region beginning at nucleotide 337 (see 053430). The
 CC sequences derived from the PCR and gt10 library enabled the
 CC compilation of the sequence in 053430.
 SQ Sequence 219 AA;
 SQ 10 A; 9 R; 11 N; 12 D; 0 B; 5 C; 10 Q; 10 E; 0 Z; 12 G; 2 H;
 SQ 8 I; 13 L; 15 K; 3 M; 8 F; 10 P; 35 S; 20 T; 4 W; 10 Y; 12 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:45-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.33
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 PEEVTRPNYL
 || ||
 DIVMSQSPSSLAVSVAEKVSMCKSSQSLFNSRTRKNYLTWYQKPGQSPKPLIYWASTRESGVDPDRFTGSG
 10 20 30 40 50 60 70
 SGTDFTLTISSVQAE DL
 80

10. US-08-249-182-5 (1-10)

R04495 HIV fusion protein PB1mn

ID R04495 standard; protein; 229 AA.
 AC R04495;
 DT 20-SEP-1990 (first entry)
 DE HIV fusion protein PB1mn
 KW HIV; fusion protein; PB1mn; therapy; AIDS; principal neutralising domain;
 KW antibodies; diagnosis; prophylaxis.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 36..203
 FT label= HIV portion of PB1mn
 PN W09003984-A.
 PD 19-APR-1990.
 PF 29-SEP-1989; U04302.
 PR 19-SEP-1989; US-407663, US-252949; WO-U04302.
 PA (REPL-) Repligen Corp.
 PI Rusche JR, Putney SD, Javaherian K, Farley J, Grimaila R, Lynn D,
 PI Petro-Breyer J.
 DR WPI; 90-147824/19.
 DR N-PSDB; 004276
 PT Principal neutralising domain of HIV variants - used for producing
 PT peptide(s) and antibodies for diagnosis; prophylaxis; and/or therapy
 PT of HIV infection.
 PS Disclosure; 108pp; English.
 CC The protein can be expressed in simian cells, and synthesis of HIV
 CC proteins can be detected immunologically. The recombinant protein
 CC product comprises a principal neutralising domain. The neutralising
 CC domain is bounded by cysteine residues which occur at positions
 CC 296 and 331, the segments between the residues form a loop.
 CC See also R04427-R04506 and 004273-004279.
 SQ Sequence 229 AA;
 SQ 9 A; 17 R; 17 N; 12 D; 0 B; 6 C; 9 Q; 13 E; 0 Z; 16 G; 4 H;
 SQ 24 I; 13 L; 17 K; 4 M; 11 F; 10 P; 13 S; 19 T; 3 W; 4 Y; 8 V;

PS Disclosure; Page 47-52; 67pp; English.
 CC The sequences given in R26982-3 contain part of the exotoxin A (ETA)
 CC sequence corresponding to positions 252-613 of the full exotoxin A
 CC sequence. These sequences are encoded by Fv(FRP5)-ETA fusion genes.
 CC The ETA sequence was used as a marker gene so that E. coli transformed
 CC with the fusion gene could be identified. The fusion genes were
 CC expressed in E. coli and the antibodies were extracted. These
 CC recombinant antibodies can be used for the qualitative and
 CC quantitative determination of c-erbB-2. This can be used for
 CC monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
 SQ Sequence 637 AA;
 SQ 69 A; 36 R; 15 N; 34 D; 0 B; 8 C; 37 Q; 36 E; 0 Z; 72 G; 9 H;
 SQ 22 I; 54 L; 21 K; 5 M; 24 F; 36 P; 48 S; 40 T; 11 W; 24 Y; 36 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:22-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNQMOTVFVGYGPTFK
 ||||| ||
 CAGPADSGDALLERNYPTGAEFLGDDVSFSTRGTQNWTVRLLQAHRLQLEERGYVVFVGYHGTFLAAQSI
 410 420 430 440 450 X 460 470
 VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
 480 490 500 510

15. US-08-249-182-9 (1-16)
 R34018 BW 835 VH.

ID R34018 standard; Protein; 115 AA.
 AC R34018;
 DT 02-AUG-1993 (first entry)
 DE BW 835 VH.
 KW Monoclonal antibody; MAb; hybridoma; lung; adenocarcinoma;
 KW mammary; ovary; prostate; polymorphic epithelial mucin; PEM.
 OS Synthetic.
 PN DE4133791-A.
 PD 15-APR-1993.
 PF 11-OCT-1991; 133791.
 PR 11-OCT-1991; DE-133791.
 PA (BEHW) BEHRINGWERKE AG.
 PI Bosslet K, Pfeleiderer P, Seemann G;
 DR WPI; 93-127068/16.
 DR N-PSDB; Q40046.
 PT New monoclonal antibody BW835 specific for tumour antigens -
 PT useful for diagnosis and treatment of tumours affecting the
 PT breasts, ovaries, prostate and lungs
 PS Disclosure; Fig 1a; 24pp; German.
 CC Monoclonal antibody BW 835 is produced by hybridoma cell line BW 835.
 CC The antibody strongly reacts with lung adenocarcinomas and human
 CC mammary-, ovary- and prostate carcinomas. It additionally reacts
 CC with polymorphic epithelial mucin (PEM) but does not react with
 CC normal human tissue.
 SQ Sequence 115 AA;
 SQ 7 A; 7 R; 6 N; 4 D; 0 B; 2 C; 5 Q; 6 E; 0 Z; 8 G; 1 H;
 SQ 4 I; 8 L; 5 K; 4 M; 4 F; 2 P; 12 S; 8 T; 4 W; 10 Y; 8 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:53-PDT using FindSeq

Initial Score = 6 Optimized Score = 6 Significance = 4.18
 Residue Identity = 37% Matches = 6 Mismatches = 10
 Gaps = 0 Conservative Substitutions = 0

PT recombinant antibodies directed to growth factor receptor C-erbB-2 -
 PT for diagnosing and treating tumours expressing C-erbB-2 e.g. breast
 PT or ovarian tumours
 PS Disclosure: Page 53-58; 67pp; English.
 CC The sequences given in R26982-3 contain part of the exotoxin A (ETA)
 CC sequence corresponding to positions 252-613 of the full exotoxin A
 CC sequence. These sequences are encoded by Fv(FRP5)-ETA fusion genes.
 CC The ETA sequence was used as a marker gene so that E. coli transformed
 CC with the fusion gene could be identified. The fusion genes were
 CC expressed in E. coli and the antibodies were extracted. These
 CC recombinant antibodies can be used for the qualitative and
 CC quantitative determination of c-erbB-2. This can be used for
 CC monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
 SQ Sequence 637 AA;
 SQ 68 A; 35 R; 12 N; 38 D; 0 B; 8 C; 35 Q; 36 E; 0 Z; 72 G; 10 H;
 SQ 26 I; 59 L; 21 K; 5 M; 16 F; 35 P; 53 S; 36 T; 11 W; 28 Y; 33 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:22-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||||| ||
CAGPADSGDALLERNYPTGA EFLG DGGDV SFSTRGTQNWTV ERLLQ AHRQL EERG YV F V G Y H G T F L E A A Q S I
    410      420      430      440      450 X      460      470

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYA Q D G E P D A R G R I R
    480      490      500      510
  
```

14. US-08-249-182-9 (1-16)
 R26982 (FRP5)-ETA fusion protein.

ID R26982 standard; Protein; 637 AA.
 AC R26982;
 DT 11-FEB-1993 (first entry)
 DE (FRP5)-ETA fusion protein.
 KW Monoclonal antibody; light chain; heavy chain; tumour; c-erbB-2;
 KW variable region; ETA.
 OS Pseudomonas aeruginosa PAK.
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= ompA_signal_peptide
 FT Peptide 22..29
 FT /label= FLAG_peptide_and_enterokinase_cleavage_site
 FT Domain 33..151
 FT /label= FRP5_heavy_chain_variable_domain
 FT Peptide 152..166
 FT /label= Linker
 FT Domain 167..274
 FT /label= FRP5_light_chain_variable_domain
 FT Protein 276..397
 FT /label= ETA_252-613
 PN EP-502812-A.
 PD 09-SEP-1992.
 PF 27-JAN-1992; 810056.
 PR 05-FEB-1991; EP-810079.
 PA (CIBA) CIBA GEIGY AG.
 PI Groner B, Hardman N, Harwerth I, Hynes NE, Wells WS;
 PI Zwickl M;
 DR WPI; 92-302096/37.
 DR N-PSDB; Q28262.
 PT Recombinant antibodies directed to growth factor receptor C-erbB-2 -
 PT for diagnosing and treating tumours expressing C-erbB-2 e.g. breast

PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 8 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 37 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQTVEVFGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAFLGDDGVSFSTRGTQNWTVRLLQAHRLLEERGYYVFGYHGTFLEAAQSI
380      390      400      410      420      430      440      X      450

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
      460      470      480      490
  
```

13. US-08-249-182-9 (1-16)

R26983 (FRP51)-ETA fusion protein.

ID R26983 standard; Protein; 637 AA.
 AC R26983;
 DT 11-FEB-1993 (first entry)
 DE (FRP51)-ETA fusion protein.
 KW Monoclonal antibody; light chain; heavy chain; tumour; c-erbB-2;
 KW variable region; ETA.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= ompA_signal_peptide
 FT Peptide 22..29
 FT /label= FLAG_peptide_and_enterokinase_cleavage_site
 FT Domain 33..152
 FT /label= FRP51_heavy_chain_variable_domain
 FT Peptide 153..167
 FT /label= Linker
 FT Domain 168..274
 FT /label= FRP5_light_chain_variable_domain
 FT Protein 276..397
 FT /label= ETA_252-613
 PN EP-502812-A.
 PD 09-SEP-1992.
 PF 27-JAN-1992; 810056.
 PR 05-FEB-1991; EP-810079.
 PA (CIBA) CIBA GEIGY AG.
 PI Groner B, Hardman N, Harwerth I, Hynes NE, Wels WS;
 PI Zwickl M;
 DR WPI; 92-302096/37.
 DR N-PSDB; 028263.

CC unpaired cysteine residues to binding moieties, (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 14 K; 6 M; 14 F; 38 P; 37 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMqTVFVGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAFLGDGGDVSFSTRGTQNWTVRLLQAHRLLEERGYVVFVGYHGTFLAAQSI
380      390      400      410      420      430      440      X      450

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
      460      470      480      490

```

12. US-08-249-182-9 (1-16)

R40102 Pseudomonas exotoxin for site-specific mutation wi

ID R40102 standard; Protein; 613 AA.
 AC R40102;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin for site-specific mutation with unpaired CYS.
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 20
 FT /note= "unpaired cysteine residue may replace Lys"
 FT Misc_difference 25
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 88
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 96
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 158
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 182
 FT /note= "unpaired cysteine residue may replace Arg"
 FT Misc_difference 188
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 192
 FT /note= "unpaired cysteine residue may replace Ser"
 FT Misc_difference 223
 FT /note= "unpaired cysteine residue may replace Lys"
 FT Misc_difference 245
 FT /note= "unpaired cysteine residue may replace Ser"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine

PT such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMØTVFVGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAFLGDDGVSFSTRGTØNWTVERLLQAHROLEERGYVVFVG YHGTFL EAAØSI
380      390      400      410      420      430      440 X 450

VFGGVRRARSØDLDAIWGRFYIAGDPALAYGYAØDØEPDARGRIR
      460      470      480      490
  
```

11. US-08-249-182-9 (1-16)

R40104 Pseudomonas exotoxin (K20C).

ID R40104 standard; Protein; 613 AA.
 AC R40104;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (K20C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 20
 FT /note= "unpaired cysteine residue replaces Lys"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their

PT /note= unpaired cysteine residue replaces Ser
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TAND-) TANDX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAFLGDDVSFSTRGTQNWTVRLLQAHRLGLEERGYVFGYHGTFLEAAQSI
380      390      400      410      420      430      440 X 450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
      460      470      480      490
  
```

10. US-08-249-182-9 (1-16)

R40105 Pseudomonas exotoxin (S25C).

ID R40105 standard; Protein; 613 AA.
 AC R40105;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (S25C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 25
 FT /note= "unpaired cysteine residue replaces Ser"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TAND-) TANDX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine

DE Pseudomonas exotoxin (S96C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 96
 FT /note= "unpaired cysteine residue replaces Ser"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAEFLGDGGDVSFSTRGTQNWTVRLLQAHRLQLEERGYPVFGYHGTFLEAAQSI
380      390      400      410      420      430      440 X 450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
      460      470      480      490
  
```

9. US-08-249-182-9 (1-16)

R40106 Pseudomonas exotoxin (S88C).

ID R40106 standard; Protein; 613 AA.
 AC R40106;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (S88C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 88

7. US-08-249-182-9 (1-16)

R40108 Pseudomonas exotoxin (S158C).

ID R40108 standard; Protein; 613 AA.
 AC R40108;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (S158C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 158
 FT /note= "unpaired cysteine residue replaces Ser"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMSTVFVGYGPTFK
                                |||||  ||
CAGPADSGDALLERNYPTGAFLGDDVSFSTRGTQNWTVRLLQAHRLQLEERGYVVFVGYHGTFLAAQSI
380      390      400      410      420      430      440      X      450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
      460      470      480      490

```

8. US-08-249-182-9 (1-16)

R40107 Pseudomonas exotoxin (S96C).

ID R40107 standard; Protein; 613 AA.
 AC R40107;

Residue Identity = 43% Matches = 7 Mismatches = 7
Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSMGTVFVVGYPGPTFK
      ||||| ||
CAGPADSGDALLERNYPTGAFLDGGDVSFSTRGTQNWTVRLLQAHRLGLEERGVVFVGYHGTFLCAAQSI
380      390      400      410      420      430      440      X      450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
      460      470      480      490
```

6. US-08-249-182-9 (1-16)

R40109 Pseudomonas exotoxin (R182C).

ID R40109 standard; Protein; 613 AA.
AC R40109;
DT 27-JAN-1994 (first entry)
DE Pseudomonas exotoxin (R182C).
KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW monoclonal antibody; ligand; cell surface; mutation;
KW steric unpaired cysteine; s.u.c.
OS Pseudomonas aeruginosa.
FH Key Location/Qualifiers
FT Misc_difference 182
FT /note= "unpaired cysteine residue replaces Arg"
PN W09315113-A.
PD 05-AUG-1993.
PF 15-JAN-1993; U00358.
PR 24-JAN-1992; US-825396.
PA (TAND-) TANOX BIOSYSTEMS INC.
PI Chang TW;
DR WPI; 93-258616/32.
PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
PT - such that conjugation of a binding mol. to the Cys blocks
PT receptor binding used as immuno:toxins for highly specific
PT targetting
PS Claim 3; Page 20-23; 30pp; English.
CC The new mutated toxin has an unpaired cysteine residue in
CC or near the cytotoxin's receptor-binding site, and retains the
CC same receptor-binding ability and cytotoxicity as the native
CC cytotoxins provided they are not conjugated with a binding mol.
CC The toxins are cross-linked through the free SH group of their
CC unpaired cysteine residues to binding mols. (including monoclonal
CC antibodies, fragments and other ligands) to form immunotoxins, and
CC these immunotoxins do not bind to the cell surface receptors which
CC are bound by the native cytotoxin. However, when the cross-linker
CC is cleaved and the binding mol. is released, the cytotoxin regains
CC its receptor-binding ability and its cytotoxicity.
SQ Sequence 613 AA;
SQ 68 A; 43 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 37 S; 27 T; 11 W; 18 Y; 37 V;
CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSMGTVFVVGYPGPTFK
      ||||| ||
CAGPADSGDALLERNYPTGAFLDGGDVSFSTRGTQNWTVRLLQAHRLGLEERGVVFVGYHGTFLCAAQSI
380      390      400      410      420      430      440      X      450
```

CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||||| ||
CAGPADSGDALLERNYPTGAFLGDGGDVSFSTRGTQNWTVRLLGAHRGLEERGYVFGVGHGTFLEAAQSI
380      390      400      410      420      430      440 X 450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQGEPDARGRIR
      460      470      480      490

```

5. US-08-249-182-9 (1-16)

R40110 Pseudomonas exotoxin (S188C).

ID R40110 standard; Protein; 613 AA.
 AC R40110;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (S188C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 188
 FT /note= "unpaired cysteine residue replaces Ser"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TAND-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02

CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.

SQ Sequence 613 AA;

SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;

SQ 26 I; 66 L; 14 K; 6 M; 14 F; 38 P; 37 S; 27 T; 11 W; 18 Y; 37 V;

CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||||| ||
CAGPADSGDALLERNYPTGA EFLG DGGDV SFSTRGTQNWTV ERLLQAH RQLEERG YV F V G Y H G T F L E A A Q S I
380      390      400      410      420      430      440 X 450

V F G G V R A R S Q D L D A I W R G F Y I A G D P A L A Y G Y A Q D Q E P D A R G R I R
      460      470      480      490
  
```

4. US-08-249-182-9 (1-16)

R40111 Pseudomonas exotoxin (S192C).

ID R40111 standard; Protein; 613 AA.

AC R40111;

DT 27-JAN-1994 (first entry)

DE Pseudomonas exotoxin (S192C).

KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;

KW target site; cytotoxin; unpaired cysteine; receptor; binding site;

KW monoclonal antibody; ligand; cell surface; mutation;

KW steric unpaired cysteine; s.u.c.

OS Pseudomonas aeruginosa.

FH Key Location/Qualifiers

FT Misc_difference 192

FT /note= "unpaired cysteine residue replaces Ser"

PN W09315113-A.

PD 05-AUG-1993.

PF 15-JAN-1993; U00358.

PR 24-JAN-1992; US-825396.

PA (TANO-) TANOX BIOSYSTEMS INC.

PI Chang TW;

DR WPI; 93-258616/32.

PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine

PT - such that conjugation of a binding mol. to the Cys blocks

PT receptor binding used as immuno:toxins for highly specific

PT targetting

PS Claim 3; Page 20-23; 30pp; English.

CC The new mutated toxin has an unpaired cysteine residue in

CC or near the cytotoxin's receptor-binding site, and retains the

CC same receptor-binding ability and cytotoxicity as the native

CC cytotoxins provided they are not conjugated with a binding mol.

CC The toxins are cross-linked through the free SH group of their

CC unpaired cysteine residues to binding mols. (including monoclonal

CC antibodies, fragments and other ligands) to form immunotoxins, and

CC these immunotoxins do not bind to the cell surface receptors which

PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting
 PS Claim 3; Page 20-23; 30pp; English.
 CC The new mutated toxin has an unpaired cysteine residue in
 CC or near the cytotoxin's receptor-binding site, and retains the
 CC same receptor-binding ability and cytotoxicity as the native
 CC cytotoxins provided they are not conjugated with a binding mol.
 CC The toxins are cross-linked through the free SH group of their
 CC unpaired cysteine residues to binding mols. (including monoclonal
 CC antibodies, fragments and other ligands) to form immunotoxins, and
 CC these immunotoxins do not bind to the cell surface receptors which
 CC are bound by the native cytotoxin. However, when the cross-linker
 CC is cleaved and the binding mol. is released, the cytotoxin regains
 CC its receptor-binding ability and its cytotoxicity.
 SQ Sequence 613 AA;
 SQ 68 A; 44 R; 21 N; 36 D; 0 B; 9 C; 27 Q; 43 E; 0 Z; 56 G; 15 H;
 SQ 26 I; 66 L; 15 K; 6 M; 14 F; 38 P; 36 S; 27 T; 11 W; 18 Y; 37 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:20-PDT using FindSeq

Initial Score = 7 Optimized Score = 7 Significance = 5.02
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                |||||  ||
CAGPADSGDALLERNYPTGAFLGDGGDVSFSTRGTQNWTVERRLLQAHRLQLEERGCVFVGYHGTFLEAAQSI
380      390      400      410      420      430      440  X  450

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDEPDARGRIR
      460      470      480      490
  
```

3. US-08-249-182-9 (1-16)

R40112 Pseudomonas exotoxin (K223C).

ID R40112 standard; Protein; 613 AA.
 AC R40112;
 DT 27-JAN-1994 (first entry)
 DE Pseudomonas exotoxin (K223C).
 KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
 KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
 KW monoclonal antibody; ligand; cell surface; mutation;
 KW steric unpaired cysteine; s.u.c.
 OS Pseudomonas aeruginosa.
 FH Key Location/Qualifiers
 FT Misc_difference 223
 FT /note= "unpaired cysteine residue replaces Lys"
 PN W09315113-A.
 PD 05-AUG-1993.
 PF 15-JAN-1993; U00358.
 PR 24-JAN-1992; US-825396.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PI Chang TW;
 DR WPI; 93-258616/32.
 PT Site-specifically mutated cytotoxin(s) with an unpaired cysteine
 PT - such that conjugation of a binding mol. to the Cys blocks
 PT receptor binding used as immuno:toxins for highly specific
 PT targetting

1. US-08-249-182-9 (1-16)

R37451 Autotaxin peptide ATX 101.

ID R37451 standard; peptide; 16 AA.
AC R37451;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 101.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 101. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 16 AA;
SQ 0 A; 0 R; 1 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 2 G; 1 H;
SQ 0 I; 0 L; 1 K; 0 M; 2 F; 1 P; 1 S; 2 T; 0 W; 1 Y; 3 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 15 Optimized Score = 15 Significance = 11.71
Residue Identity = 93% Matches = 15 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSM0TVFVGYPGTFK
||| |||||
VNSH0TVFVGYPGTFK
X      10      X

```

2. US-08-249-182-9 (1-16)

R40113 Pseudomonas exotoxin (S245C).

ID R40113 standard; Protein; 613 AA.
AC R40113;
DT 27-JAN-1994 (first entry)
DE Pseudomonas exotoxin (S245C).
KW Pseudomonas exotoxin; PE; diphtheria toxin; DT; immunotoxin;
KW target site; cytotoxin; unpaired cysteine; receptor; binding site;
KW monoclonal antibody; ligand; cell surface; mutation;
KW steric unpaired cysteine; s.u.c.
OS Pseudomonas aeruginosa.
FH Key Location/Qualifiers
FT Misc_difference 245
FT /note= "unpaired cysteine residue replaces Ser"
PN W09315113-A.
PD 05-AUG-1993.

Scores: Mean Median Standard Deviation
 1 3 1.20

Times: CPU Total Elapsed
 00:00:08.92 00:00:08.00

Number of residues: 482836
 Number of sequences searched: 5543
 Number of scores above cutoff: 3295

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

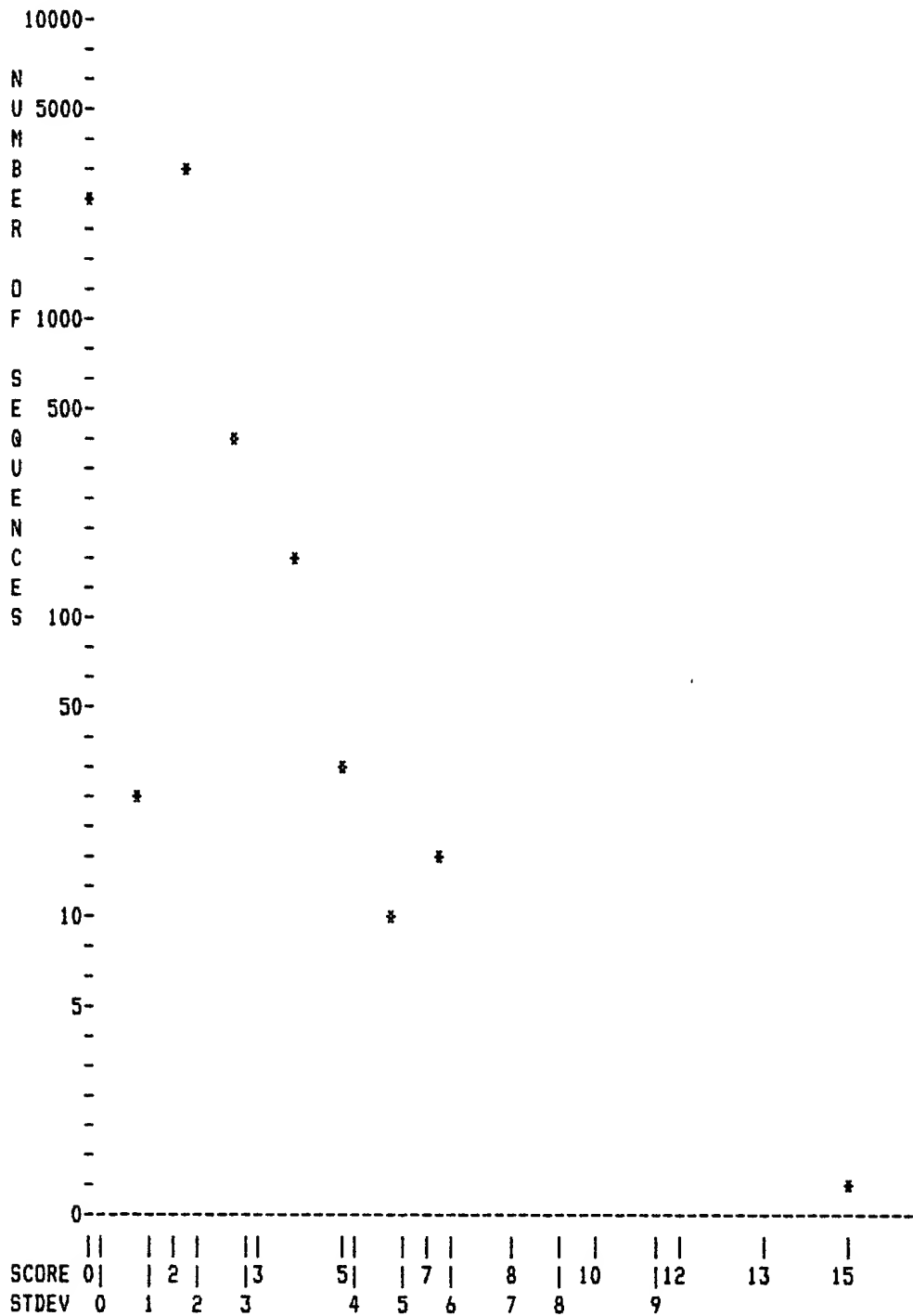
A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 11 standard deviations above mean ****						
1. R37451	Autotaxin peptide ATX 101.	16	15	15	11.71	0
**** 5 standard deviations above mean ****						
2. R40113	Pseudomonas exotoxin (S245C).	613	7	7	5.02	0
3. R40112	Pseudomonas exotoxin (K223C).	613	7	7	5.02	0
4. R40111	Pseudomonas exotoxin (S192C).	613	7	7	5.02	0
5. R40110	Pseudomonas exotoxin (S188C).	613	7	7	5.02	0
6. R40109	Pseudomonas exotoxin (R182C).	613	7	7	5.02	0
7. R40108	Pseudomonas exotoxin (S158C).	613	7	7	5.02	0
8. R40107	Pseudomonas exotoxin (S96C).	613	7	7	5.02	0
9. R40106	Pseudomonas exotoxin (S88C).	613	7	7	5.02	0
10. R40105	Pseudomonas exotoxin (S25C).	613	7	7	5.02	0
11. R40104	Pseudomonas exotoxin (K20C).	613	7	7	5.02	0
12. R40102	Pseudomonas exotoxin for site	613	7	7	5.02	0
13. R26983	(FRP51)-ETA fusion protein.	637	7	7	5.02	0
14. R26982	(FRP5)-ETA fusion protein.	637	7	7	5.02	0
**** 4 standard deviations above mean ****						
15. R34018	BW 835 VH.	115	6	6	4.18	0
16. R09423	Br-3 Heavy Chain V Region (mo	143	6	6	4.18	0
17. R28582	HCV amino acid sequence contg	2436	6	6	4.18	0
18. R29527	HCV antigen T7N1-30.	2510	6	6	4.18	0
19. R24440	Composite HCV HC-J1/CDC/CHI p	2894	6	6	4.18	0
20. R20111	Non-A, non-B viral genome pro	3010	6	6	4.18	0
21. R20091	Non-A, non-B viral genome pro	3010	6	6	4.18	0
22. R31621	Hepatitis C virus (HCV) polyp	3011	6	6	4.18	0
23. R22154	NANBV Hutch c59 isolate genom	3011	6	6	4.18	0
**** 3 standard deviations above mean ****						
24. R33087	Human cytomegalovirus antibod	31	5	5	3.35	0
25. R33086	Human cytomegalovirus antibod	31	5	5	3.35	0
26. R33089	Human cytomegalovirus antibod	32	5	5	3.35	0
27. R33091	Human cytomegalovirus antibod	51	5	5	3.35	0
28. R25410	Heavy chain variable domain o	114	5	5	3.35	0
29. R40215	Sequence of mouse hybridoma c	119	5	5	3.35	0
30. R12358	Heavy chain variable region o	134	5	5	3.35	0
31. R12326	Heavy chain variable region o	136	5	5	3.35	0
32. R42212	CRABP-I gene product.	137	5	5	3.35	0
33. R30484	VH region of Ab to pre-S2 ant	139	5	5	3.35	0
34. R32128	Anti-IL2R-alpha antibody M-21	183	5	5	3.35	0
35. R15326	IL-2 chimeric antibody heavy	183	5	5	3.35	0
36. R33252	HBsAg encoded by pGPD-1(HBS).	226	5	5	3.35	0
37. R11496	RP142/HBsAg.	250	5	5	3.35	0
38. R11495	RP135/HBsAg.	251	5	6	3.35	0

Query sequence being compared:US-08-249-182-9 (1-16)
 Number of sequences searched: 5543
 Number of scores above cutoff: 3295

Results of the initial comparison of US-08-249-182-9 (1-16) with:
 File : /home/shears/loring/lorin*.pep



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

GNLACLSLCHIETERAPSRAPTITLKKTPKPKTTKKPTKTTIHHRTSPETKLQPKNNTATPQQGILSSTEHH
180 190 200 210 220 230 240

TNQTSTQI

250

15. US-08-249-182-8 (1-11)

R29637 pCTD ORF 2.

ID R29637 standard; Protein; 354 AA.
AC R29637;
DT 05-FEB-1993 (first entry)
DE pCTD ORF 2.
KW CT; pCTD; epithelium; ocula mucosa; uro-genital mucosa; antigen;
KW monoclonal; polyclonal; antibody; vaccine.
OS Chlamydia trachomatis.
PN EP-499681-A.
PD 26-AUG-1992.
PF 17-APR-1991; 106110.
PR 07-FEB-1991; IT-000314.
PA (ISTS) SCLAVO SPA.
PI Comanducci M, Giuliani MM, Ratti G, Tecce MF;
DR WPI; 92-285922/35.
DR N-PSDB; Q27429.
PT PCTD plasmid from Chlamydia Trachomatis and immunogenic proteins
PT - for diagnosing and vaccinating against Chlamydia infections
PT e.g. venereal lymphogranuloma
PS Claim 1; Page 8-16; 40pp; English.
CC The sequences given in R29636-43 are encoded by the plasmid isolated
CC from Chlamydia trachomatis (CT) serotype D, pCTD. This serotype
CC generally infects epithelial tissues, such as the ocular and
CC uro-genital mucous membranes, and shows a low virulence. Of the eight
CC proteins encoded by the plasmid, seven are encoded by the sense strand
CC and the eighth is encoded by the complementary strand. These proteins
CC can be used as antigens for the preparation of poly- and mono-clonal
CC antibodies to be used in diagnostics. The antigens can also be used
CC in the formulation of vaccines against infections due to CT.
SQ Sequence 354 AA;
SQ 20 A; 23 R; 21 N; 10 D; 0 B; 4 C; 14 Q; 29 E; 0 Z; 14 G; 7 H;
SQ 32 I; 31 L; 32 K; 5 M; 16 F; 14 P; 23 S; 20 T; 7 W; 16 Y; 16 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:22-PDT using FindSeq

Initial Score = 5 Optimized Score = 6 Significance = 3.30
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10
GGQPLWITATK
I I I I
FQFEGWIPRIRFTKTEFLEAYGVKRYKTSRNKYEFSGKEAETALEALYHLGHQPFLLIVATRTWNGTQIVD
90 100 110 120 130 X 140 X 150

RYQTLSPRIIRIYEGWEALTDEENIDIDLTPFNSPPTRKH
160 170 180 190

> 0 <
0| 10 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-08-249-182-9.res made by on Wed 21 Sep 94 12:04:15-PDT.

Initial Score = 5 Optimized Score = 6 Significance = 3.30
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                X      X
              GGQPLWITATK
                || | || |
MSDQATTLRIKPLGDRILVKREEDSTARGGIILPDTAKKKQDRAEVLVLGTGKRDKDGNLLPFVTVGDTV
      10      20      30      40      50      60      70

LIDKYAGQELTVDGEEYV
      80      90
  
```

14. US-08-249-182-8 (1-11)
 R21326 Sequence of protein G.

ID R21326 standard; Protein; 257 AA.
 AC R21326;
 DT 17-MAY-1992 (first entry)
 DE Sequence of protein G.
 KW Subunit vaccine; recombinant virus; antibody; probe;
 KW passive immunisation.
 OS Bovine respiratory syncytium virus isolate 391-2.
 FH Key Location/Qualifiers
 FT Domain 1..37
 FT /label= cytoplasmic
 FT Domain 38..65
 FT /label= transmembrane
 FT Domain 66..257
 FT /label= extracellular
 PN W09201471-A.
 PD 06-FEB-1992.
 PF 23-JUL-1991; U05194.
 PR 24-JUL-1990; US-557267.
 PA (UABR-) UAB RES FOUND.
 PI Wertz GW, Lerch R;
 DR WPI; 92-064708/08.
 DR N-PSDB; Q21150.
 PT DNA encoding bovine respiratory syncytium virus - G,F and N
 PT proteins, used to produce recombinant proteins and antibodies,
 PT for vaccine and diagnosis of BRS viral infection
 PS Claim 22; Fig 3; 124pp; English.
 CC The inventors claim recombinant DNA mols. which comprise nucleic
 CC acid sequences encoding bovine respiratory syncytial (BRS) virus
 CC proteins G, F and N respectively, and BRS virus G, F and N proteins
 CC and fragments prep'd. by culturing transformed microorganisms and
 CC cells. G-protein DNA is pref. contained in plasmid pRLG414-76-191
 CC (ATCC NO 40841), F-protein DNA in pRLF2012-76-1902 (ATCC No 40842),
 CC N-protein in pRLNB3-76 (ATCC No 40843). For expression, a
 CC regulatory sequence is also included. The resulting mols. are
 CC contained in rVG642 (ATCC No VR2276), and rVF-464 (ATCC Vr2277) for
 CC G-protein and F-protein respectively.
 SQ Sequence 257 AA;
 SQ 12 A; 7 R; 16 N; 2 D; 0 B; 5 C; 17 Q; 8 E; 0 Z; 8 G; 14 H;
 SQ 18 I; 21 L; 23 K; 2 M; 6 F; 20 P; 24 S; 42 T; 1 W; 6 Y; 5 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:50-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

KW GroEL.
 OS Chlamydia trachomatis serovar A.
 PN US7531317-A.
 PD 09-JUL-1991.
 PF 31-MAY-1990; 143560.
 PR 31-MAY-1990; US-531317.
 PA (USSH) NAT INST OF HEALTH.
 DR WPI; 91-245693/33.
 DR N-PSDB; Q13137.
 PT DNA encoding HypA and HypB Chlamydia proteins - used to develop
 PT prods. for detection of and vaccines against Chlamydia infection.
 PS Disclosure; Fig 7; 51pp; English.
 CC The sequence was deduced from the first of two ORFs found in the
 CC hyp operon of clone pTA571CC prepd. from C.trachomatis genomic DNA.
 CC It is the HypA hypersensitivity protein, analogous to the E. coli
 CC GroES protein. It can be used to to raise antibodies and to
 CC prepare vaccines for the treatment of Chlamydial infections.
 CC See also R13334-R13337.
 SQ Sequence 102 AA;
 SQ 7 A; 4 R; 0 N; 8 D; 0 B; 0 C; 8 Q; 10 E; 0 Z; 9 G; 0 H;
 SQ 8 I; 10 L; 10 K; 2 M; 1 F; 3 P; 4 S; 6 T; 0 W; 2 Y; 10 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:33-PDT using FindSeq

Initial Score = 5 Optimized Score = 6 Significance = 3.30
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                X      X
              GGQPLWITAK
                || | || |
MSDQATTLKIKPLGDRILVKREEEASTARGGIILPDTAKKKQDRAEVLALGTGKKDDKGQQLPFEVQVGDIV
      10      20      30      40      50      60      70

LIDKYSGQELTVEGEEYV
      80      90
  
```

13. US-08-249-182-8 (1-11)
 R13334 HypA protein.

ID R13334 standard; Protein; 102 AA.
 AC R13334;
 DT 22-OCT-1991 (first entry)
 DE HypA protein.
 KW Antibodies; heat shock; hypersensitive; allergen; HSP60; GroES.
 OS Chlamydia psittaci GPIC.
 PN US7531317-A.
 PD 09-JUL-1991.
 PF 31-MAY-1990; 143560.
 PR 31-MAY-1990; US-531317.
 PA (USSH) NAT INST OF HEALTH.
 DR WPI; 91-245693/33.
 DR N-PSDB; Q13136.
 PT DNA encoding HypA and HypB Chlamydia proteins - used to develop
 PT prods. for detection of and vaccines against Chlamydia infection.
 PS Disclosure; Fig 5; 51pp; English.
 CC The sequence was deduced from the first of two ORFs found in clone
 CC pGP57, prepd. from C. psittaci genomic DNA and contg. the Hyp
 CC operon. It is the hypA protein, of approx. 12 kD, analogous to
 CC the GroES heat shock protein of E. coli. The recombinant protein
 CC can be used to to raise antibodies and in the preparation of
 CC vaccines for the treatment of Chlamydial infections.
 CC See also R13335-R13337.
 SQ Sequence 102 AA;
 SQ 6 A; 6 R; 1 N; 10 D; 0 B; 0 C; 4 Q; 10 E; 0 Z; 9 G; 0 H;

CC antibodies by immunoassay methods.
SQ Sequence 70 AA;
SQ 6 A; 4 R; 1 N; 2 D; 0 B; 2 C; 1 Q; 4 E; 0 Z; 4 G; 2 H;
SQ 1 I; 7 L; 5 K; 4 M; 1 F; 1 P; 2 S; 6 T; 7 W; 4 Y; 6 V;
CC Retrieved by shears on Wed 21 Sep 94 11:56:53-PDT using FindSeq

Initial Score = 5 Optimized Score = 6 Significance = 3.30
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X

X

GGQPLWITATK

I I I I I

MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVRVYGVPLWKEATTTLFCASDAKAYDTEVHNVWA

10203040506070

11. US-08-249-182-8 (1-11)
P91962 Polypeptide encoded by cDNA 50F1.

ID P91962 standard; polypeptide; 74 AA.
AC P91962;
DT 21-FEB-1990 (first entry)
DE Polypeptide encoded by cDNA 50F1.
KW Human colonic mucosa; colonic cancer; antibody.
OS Homo sapiens.
PN EP-337498-A.
PD 18-OCT-1989.
PF 17-APR-1989; 106875.
PR 15-APR-1988; US-182185.
PA (MONT-) Montefiore Medical Center.
PI Augenlicht LH;
DR WPI; 89-302259/42.
DR N-PSDB; N91510.
PT Monitoring stage of malignant disease, esp. colonic cancer - by measuring
PT relative abundance of specific RNA, also new therapeutic RNA, polypeptide
PT and antibody.
PS Disclosure; Fig. 7; 20pp; English.
CC The polypeptide is expressed by cDNA 50F1. It can be used treating and
CC preventing colonic cancer of human colonic mucosa. Mono- or polyclonal
CC antibodies raised against the polypeptide can be used for its purificn.
CC and for immunoassay of the polypeptide.
SQ Sequence 74 AA;
SQ 1 A; 8 R; 4 N; 1 D; 0 B; 3 C; 2 Q; 2 E; 0 Z; 10 G; 6 H;
SQ 7 I; 8 L; 1 K; 1 M; 4 F; 1 P; 6 S; 6 T; 0 W; 0 Y; 3 V;
CC Retrieved by shears on Wed 21 Sep 94 11:56:51-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
Residue Identity = 45% Matches = 5 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X

10

GGQPLWITATK

I I I I I

GGGGGGGGHTSITAHHSLIENNRNQMRFKHCCITIVTGSFLTPTSLRVLRVSLHIDRRHLRLNIF

X 10X 2030405060

12. US-08-249-182-8 (1-11)
R13336 HypA protein.

ID R13336 standard; Protein; 102 AA.
AC R13336;
DT 22-OCT-1991 (first entry)
DE HypA protein.

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
          || ||
ZRDNWRSELYKYKVIKIEPLGIAPTKAKRRVVQREKRY
      10      X 20      X 30

```

10. US-08-249-182-8 (1-11)

P94662 Protein sequence for the amino terminal portion of

ID P94662 standard; protein; 70 AA.
 AC P94662;
 DT 28-JUN-1990 (first entry)
 DE Protein sequence for the amino terminal portion of the HTLV-III
 DE envelope gp 160 and gp 120
 KW Glycoproteins gp120 and gp160; HTLV-III antibodies; immunoassay;
 KW HTLV-III.
 OS HTLV-III.
 FH Key Location/Qualifiers
 FT Cleavage 90..91
 FT /note="site at which leader sequence cleaved from gp 160"
 FT Region 34
 FT /note="residue determined by radiolabel sequence
 FT analysis"
 FT Region 36
 FT /note="as above"
 FT Region 38
 FT /note="as above"
 FT Region 42
 FT /note="as above"
 FT Region 44
 FT /note="as above"
 FT Region 52
 FT /note="as above"
 FT Region 54
 FT /note="as above"
 FT Region 65
 FT /note="as above"
 FT Region 68
 FT /note="as above"
 PN CA1247082-A.
 PD 20-DEC-1988.
 PF 12-NOV-1985; 495112.
 PR 07-NOV-1985; US-795974, US-670361.
 PA (HARD) Harvard College.
 PI Essex ME;
 DR WPI; 89-061499/09.
 DR N-PSDB; N94662.
 PT Human T-cell lymphotropic virus type III -
 PT useful for detection of human T-cell lymphotropic virus type
 PT III antibodies.
 PS Disclosure; ; 27pp; English.
 CC New polypeptides are immunologically cross-reactive with glycoproteins
 CC gp 120 or gp 160 present in cells infected with HTLV-III, where gp 120
 CC has a m.wt. of ca. 120,000 and gp 160 has a m.wt. of ca. 160,000
 CC (90,000 inunglycosylated form). Also claimed are a method of assaying
 CC a biological specimen for HTLV-III antibodies, comprising incubating the
 CC specimen with the polypeptides, and a kit for this use. The polypeptides
 CC comprise gp 120, gp 160 and their unglycosylated forms opt. labelled or
 CC bound to an insoluble phase, or anti-idiotypic antibodies having an
 CC antigenic determinant cross-reactive with gp120. Gp 160 is a precursor

PI Avrameas A, Guillet JG, Moraillon A, Strosberg AD;
 DR WPI; 92-217024/26.
 PT Envelope protein fragments of feline immunodeficiency virus and
 PT derived antibodies - useful for diagnosis, therapy and in
 PT construction of model for HIV
 PS Claim 8; Page 30; 47pp; French.
 CC The peptide is derived from Petaluma strain FIV. It is useful as an
 CC immunological reagent for the detection, diagnosis and monitoring of
 CC immune deficiency in cats. It is recognised by at least one antibody
 CC produced during infection and/or inoculation by FIV Petaluma strain.
 CC For therapeutic use, the antibodies are conjugated with an oxidative
 CC enzyme.
 SQ Sequence 28 AA;
 SQ 0 A; 0 R; 2 N; 0 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 3 G; 0 H;
 SQ 2 I; 5 L; 4 K; 1 M; 0 F; 0 P; 2 S; 2 T; 0 W; 2 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:17-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || |  ||
LQGKINISLCLTGGKMLYNKVTQQLSYC
    10 X      20 X

```

9. US-08-249-182-8 (1-11)

P91359 Amino acids 482-517 of HIV glycoprotein gp120

ID P91359 standard; peptide; 38 AA.
 AC P91359;
 DT 12-APR-1990 (first entry)
 DE Amino acids 482-517 of HIV glycoprotein gp120
 KW Human immunodeficiency virus; HIV; glycoprotein; gp120;
 KW immunoassay reagent; HIV antibodies; HIV vaccine.
 OS Human immunodeficiency virus (HIV).
 PN EP-317804-A.
 PD 31-MAY-1989.
 PF 03-NOV-1988; 118299.
 PR 24-NOV-1987; US-124801.
 PA (ABBQ) Abbott Laboratories.
 PI Sarin VK, Knigge KM;
 DR WPI; 89-158923/22.
 PT New HIV peptide(s)
 PT - useful as immunoassay reagents, for prodn. of
 PT antibodies and in vaccines
 PS Claim 1(b); 14pp; English.
 CC It is called Peptide II. It is encoded in the region between
 CC bp 7197-7305. Z = -H, a blocking group, an amino acid, a peptide, protein
 CC or any linker. X = -OH, -NH2, -NR(1)R(2) (wherein R corresponds to an
 CC alkyl group), a peptide, protein or any linker. It may be produced by
 CC a number of methods including solid phase synthesis and recombinant DNA
 CC methodology. It provides a non-infectious and pure source of HIV antigens
 CC for use as immunoassay reagents for detection of HIV antibodies in
 CC biological samples, esp. human saliva, for prodn. of poly- and
 CC monoclonal HIV antibodies, and in HIV vaccines. It can be used as a
 CC source of HIV antigens alone or in combination or may be linked to
 CC larger carrier molecules.
 SQ Sequence 38 AA;
 SQ 2 A; 6 R; 1 N; 1 D; 0 B; 0 C; 1 Q; 3 E; 0 Z; 1 G; 0 H;
 SQ 3 I; 2 L; 6 K; 0 M; 0 F; 2 P; 1 S; 1 T; 1 W; 2 Y; 3 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:56:49-PDT using FindSeq

7. US-08-249-182-8 (1-11)

P98460 Sequence of C. trachomatis serovar J major outer m

ID P98460 standard; Protein: 14 AA.
 AC P98460;
 DT 06-MAR-1992 (first entry)
 DE Sequence of C. trachomatis serovar J major outer membrane protein (MOMP)
 DE variable domain (VD) J-VDIII encoded by base pairs 742-783
 KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay.
 OS Chlamydia trachomatis.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97095.
 PT Chlamydia trachomatis genes - used for determn. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 17; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 14 AA;
 SQ 4 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:43-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

X      X
GGGPLWITATK
  || |||
AEFPLDITAGTEAA
X      10

```

8. US-08-249-182-8 (1-11)

R24868 Sequence of peptide fragment V which corresp. to r

ID R24868 standard; Protein: 28 AA.
 AC R24868;
 DT 03-JAN-1992 (first entry)
 DE Sequence of peptide fragment V which corresp. to residues 292-320 of
 DE the env protein.
 KW env protein; immunological reagent; diagnosis; therapy; antibody.
 OS Feline immunodeficiency virus Petaluma strain.
 PN W09209632-A.
 PD 11-JUN-1992.
 PF 20-NOV-1991; F00917.
 PR 21-NOV-1990; FR-014519.
 PA (USPHS) US DEPT HEALTH & HUMAN.
 PI (USPHS) US DEPT HEALTH & HUMAN.
 DR WPI; 89-339697/46.
 DR N-PSDB; N97095.
 PT Chlamydia trachomatis genes - used for determn. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 17; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.

CC methodologies applicable in the development of new diagnostic tests
CC for serotyping.
SQ Sequence 14 AA;
SQ 4 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:43-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
Residue Identity = 45% Matches = 5 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X X
GGGPLWITATK
|| |||
AEFPLNITAGTEAA
X 10

6. US-08-249-182-8 (1-11)

P98468 Sequence of *C. trachomatis* serovar L3 major outer

ID P98468 standard; Protein; 14 AA.
AC P98468;
DT 06-MAR-1992 (first entry)
DE Sequence of *C. trachomatis* serovar L3 major outer membrane protein (MOMP)
DE variable domain (VD) L3-VDIII encoded by base pairs 742-783
KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
KW diagnosis; serotyping; non-immunologic assay.
OS *Chlamydia trachomatis*.
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US DEPT HEALTH & HUMAN.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
DR WPI; 89-339697/46.
DR N-PSDB; N97103.
PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
PT amino sequences of the variable domains of the major outer
PT membrane proteins
PS Disclosure; Fig 19; 49pp; English.
CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
CC serovars and the amino acid sequences were deduced. The MOMP VDs
CC with the greatest total hydrophilicity and charge values were found
CC to be the location of antigenic determinants recognised by MOMP
CC specific monoclonal antibodies. The nucleotide, amino acid
CC sequences and hydrophilicity/charge value analyses will assist in
CC the selection of appropriate MOMP antigenic determinants to be used
CC in the construction of synthetic peptides, subunits or recombinant
CC chlamydial vaccines. This will allow the prodn. of reagents and
CC methodologies applicable in the development of new diagnostic tests
CC for serotyping.
SQ Sequence 14 AA;
SQ 4 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:43-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
Residue Identity = 45% Matches = 5 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X X
GGGPLWITATK
|| |||
AEFPLDITAGTEAA
X 10

PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI: 89-339697/46.
 DR N-PSDB: N97083.
 PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure: Fig 14; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 14 AA;
 SQ 4 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:43-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
    GGQPLWITATK
      || |||
    AEFPLDITAGTEAA
      X      10
  
```

5. US-08-249-182-8 (1-11)

P98444 Sequence of *C. trachomatis* serovar C major outer m

ID P98444 standard; Protein; 14 AA.
 AC P98444;
 DT 06-MAR-1992 (first entry)
 DE Sequence of *C. trachomatis* serovar C major outer membrane protein (MOMP)
 DE variable domain (VD) C-VDIII encoded by base pairs 742-783
 KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay;
 OS *Chlamydia trachomatis*.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI: 89-339697/46.
 DR N-PSDB: N97079.
 PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure: Fig 13; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and

P98452 Sequence of C. trachomatis serovar H major outer m

ID P98452 standard; Protein; 14 AA.
AC P98452;
DT 06-MAR-1992 (first entry)
DE Sequence of C. trachomatis serovar H major outer membrane protein (MOMP)
DE variable domain (VD) H-VDIII encoded by base pairs 742-783
KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
KW diagnosis; serotyping; non-immunologic assay; ss.
OS Chlamydia trachomatis.
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US DEPT HEALTH & HUMAN.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
DR WPI; 89-339697/46.
DR N-PSDB; N97087.
PT Chlamydia trachomatis genes - used for determn. of nucleotide and
PT amino sequences of the variable domains of the major outer
PT membrane proteins
PS Disclosure; Fig 15; 49pp; English.
CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
CC serovars and the amino acid sequences were deduced. The MOMP VDs
CC with the greatest total hydrophilicity and charge values were found
CC to be the location of antigenic determinants recognised by MOMP
CC specific monoclonal antibodies. The nucleotide, amino acid
CC sequences and hydrophilicity/charge value analyses will assist in
CC the selection of appropriate MOMP antigenic determinants to be used
CC in the construction of synthetic peptides, subunits or recombinant
CC chlamydial vaccines. This will allow the prodn. of reagents and
CC methodologies applicable in the development of new diagnostic tests
CC for serotyping.
SQ Sequence 14 AA;
SQ 4 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:43-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
Residue Identity = 45% Matches = 5 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X X
GGGPLWITATK
|| |||
AEFPLDITAGTEAA
X 10

4. US-08-249-182-8 (1-11)

P98448 Sequence of C. trachomatis serovar A major outer m

ID P98448 standard; Protein; 14 AA.
AC P98448;
DT 06-MAR-1992 (first entry)
DE Sequence of C. trachomatis serovar A major outer membrane protein (MOMP)
DE variable domain (VD) A-VDIII encoded by base pairs 742-783
KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
KW diagnosis; serotyping; non-immunologic assay.
OS Chlamydia trachomatis.
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US DEPT HEALTH & HUMAN.

Sequence 11 AA;
SQ 1 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 2 G; 0 H;
SQ 1 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 2 T; 1 W; 0 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 11 Optimized Score = 11 Significance = 8.24
Residue Identity = 100% Matches = 11 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X X
GGGPLWITATK
|||||
GGGPLWITATK
X 10

2. US-08-249-182-8 (1-11)

P98464 Sequence of *C. trachomatis* serovar K major outer m

ID P98464 standard; Protein; 14 AA.
AC P98464;
DT 06-MAR-1992 (first entry)
DE Sequence of *C. trachomatis* serovar K major outer membrane protein (MOMP)
DE variable domain (VD) K-VDIII encoded by base pairs 742-783
KW *Chlamydia trachomatis*; antigen; monoclonal antibody; vaccine;
KW diagnosis; serotyping; non-immunologic assay.
OS *Chlamydia trachomatis*.
PN US7324664-A.
PD 29-AUG-1989.
PF 17-MAR-1989; 324664.
PR 17-MAR-1989; US-324664.
PA (USSH) US DEPT HEALTH & HUMAN.
PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
DR WPI; 89-339697/46.
DR N-PSDB; N97099.
PT *Chlamydia trachomatis* genes - used for determ. of nucleotide and
PT amino sequences of the variable domains of the major outer
PT membrane proteins
PS Disclosure; Fig 18; 49pp; English.
CC The inventors sequenced the 4 MOMP VDs of ten *C. trachomatis*
CC serovars and the amino acid sequences were deduced. The MOMP VDs
CC with the greatest total hydrophilicity and charge values were found
CC to be the location of antigenic determinants recognised by MOMP
CC specific monoclonal antibodies. The nucleotide, amino acid
CC sequences and hydrophilicity/charge value analyses will assist in
CC the selection of appropriate MOMP antigenic determinants to be used
CC in the construction of synthetic peptides, subunits or recombinant
CC chlamydial vaccines. This will allow the prodn. of reagents and
CC methodologies applicable in the development of new diagnostic tests
CC for serotyping.
SQ Sequence 14 AA;
SQ 3 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:44-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.30
Residue Identity = 45% Matches = 5 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

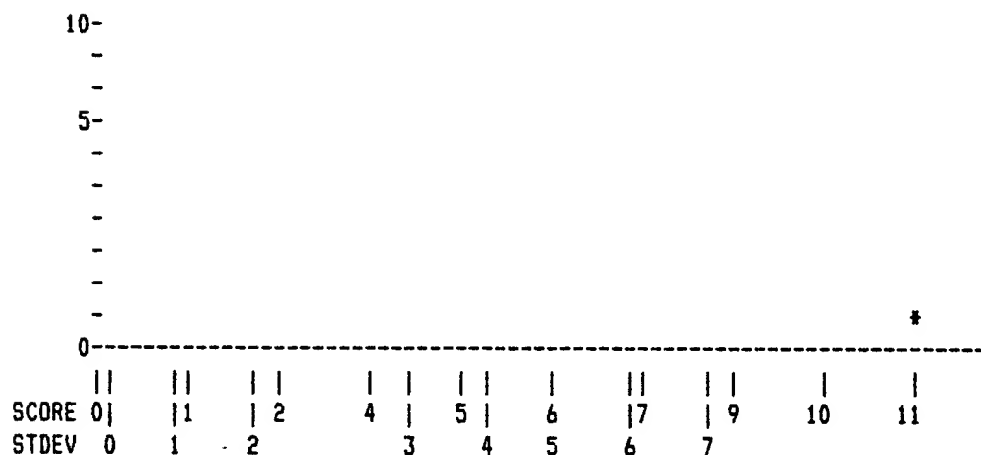
X X
GGGPLWITATK
|||
VEFPLDITAGTEAA
X 10

7. P98460	Sequence of C. trachomatis se	14	5	5	3.30	0
8. R24868	Sequence of peptide fragment	28	5	5	3.30	0
9. P91359	Amino acids 482-517 of HIV gl	38	5	5	3.30	0
10. P94662	Protein sequence for the amin	70	5	6	3.30	0
11. P91962	Polypeptide encoded by cDNA 5	74	5	5	3.30	0
12. R13336	HypA protein.	102	5	6	3.30	0
13. R13334	HypA protein.	102	5	6	3.30	0
14. R21326	Sequence of protein G.	257	5	5	3.30	0
15. R29637	pCTD DRF 2.	354	5	6	3.30	0
16. R39684	VCAM-6D/ICAM-1.	643	5	5	3.30	0
17. R39685	VCAM-6D/ICAM-2.	644	5	5	3.30	0
18. R39686	VCAM-6D/ICAM4-1.	647	5	5	3.30	0
19. R38549	VCAM-6D.	647	5	5	3.30	0
20. R05795	HIV-1 env mutein lacking hype	700	5	5	3.30	0
21. R38550	VCAM/ICAM-1.	735	5	5	3.30	0
22. R39682	VCAM/ICAM-2.	736	5	5	3.30	0
23. R38548	VCAM-7D.	739	5	5	3.30	0
24. R08118	Vascular cell adhesion molecu	739	5	5	3.30	0
25. R39683	VCAM/ICAM-3.	740	5	5	3.30	0
26. R29706	env gene decoded from viral D	863	5	5	3.30	0
27. R43950	Ge protein fragment.	1150	5	5	3.30	0
28. R12608	EGRF-R erbB-3 clone E3-16 pro	1343	5	5	3.30	0
**** 2 standard deviations above mean ****						
29. P93053	HIV env protein analogue (j).	11	4	4	2.47	0
30. R42333	EBV VCA peptide.	12	4	5	2.47	0
31. R42332	EBV VCA peptide.	12	4	4	2.47	0
32. R42331	EBV VCA peptide.	12	4	4	2.47	0
33. R42330	EBV VCA peptide.	12	4	4	2.47	0
34. R41295	Peptide fragment F7.	12	4	5	2.47	0
35. P98436	Sequence of C. trachomatis se	14	4	4	2.47	0
36. P98428	Sequence of C. trachomatis se	14	4	4	2.47	0
37. P98420	Sequence of C. trachomatis se	14	4	4	2.47	0
38. P98412	Sequence of C. trachomatis se	14	4	4	2.47	0
39. P98440	Sequence of C. trachomatis se	14	4	4	2.47	0
40. P93291	Sequence of Chlamydia trachom	14	4	4	2.47	0

1. US-08-249-182-8 (1-11)

R37450 Autotaxin peptide ATX 100.

ID R37450 standard; peptide; 11 AA.
AC R37450;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 100.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 100. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.



PARAMETERS

```

Similarity matrix      Unitary      K-tuple      2
Mismatch penalty      1      Joining penalty      20
Gap penalty            1.00      Window size      5
Gap size penalty      0.05
Cutoff score          0
Randomization group   0

Initial scores to save      40      Alignments to save      15
Optimized scores to save    0      Display context      50
  
```

SEARCH STATISTICS

```

Scores:                Mean      Median      Standard Deviation
                        1          3          1.21

Times:                 CPU          Total Elapsed
                        00:00:08.92      00:00:08.00

Number of residues:      482836
Number of sequences searched: 5543
Number of scores above cutoff: 3132
  
```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.			Sig. Frame
		Length	Score	Score	
1. R37450	Autotaxin peptide ATX 100.	11	11	11	8.24 0

The list of other best scores is:

Sequence Name	Description	Init. Opt.			Sig. Frame
		Length	Score	Score	

**** 3 standard deviations above mean ****					
2. P98464	Sequence of C. trachomatis se	14	5	5	3.30 0
3. P98452	Sequence of C. trachomatis se	14	5	5	3.30 0
4. P98448	Sequence of C. trachomatis se	14	5	5	3.30 0
5. P98444	Sequence of C. trachomatis se	14	5	5	3.30 0
6. P98468	Sequence of C. trachomatis se	14	5	5	3.30 0

CC encoding one or more epitopes from the gp48 and gp21 regions of HTLV-
 CC 1 env protein. This protein and fusion proteins comprising it are
 CC useful for detecting antibodies to HTLV-1 in body fluids, eg blood,
 CC where they provide a more sensitive and selective assay than
 CC current viral lysate tests.
 CC See also Q11553-58.
 SQ Sequence 106 AA;
 SQ 5 A; 6 R; 7 N; 5 D; 0 B; 2 C; 9 Q; 5 E; 0 Z; 8 G; 3 H;
 SQ 6 I; 17 L; 6 K; 1 M; 2 F; 5 P; 7 S; 3 T; 3 W; 1 Y; 5 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:21-PDT using FindSeq

Initial Score = 4 Optimized Score = 5 Significance = 2.68
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 |||||
 KNLLKIAQYAAQNRRLDLLFWEGGLCKALQEQCRFPNITNSHVPILQERPPLENRVLTGWGLNWDLGSD
 30 40 50 60 70 80 X 90

LQPSLIS
 100
 > D <
 0| 10 IntelliGenetics
 > D <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-08-249-182-8.res made by on Wed 21 Sep 94 12:12:21-PDT.

Query sequence being compared:US-08-249-182-8 (1-11)
 Number of sequences searched: 5543
 Number of scores above cutoff: 3132

Results of the initial comparison of US-08-249-182-8 (1-11) with:
 File : /hone/shears/loring/lorin*.pep

10000-
 -
 N -
 U 5000-
 M -
 B -
 E * *
 R -
 -
 D -
 F 1000-
 -
 S -
 E 500- *
 Q -
 U -
 E -
 N -
 C - *
 E -
 S 100-
 -
 -
 50-
 -
 -

PR 17-SEP-1981;
 PR 18-SEP-1980; GB-030208.
 PR 22-OCT-1980; GB-034130.
 PR 27-NOV-1980; GB-038147.
 PR 08-APR-1981; GB-011064.
 PR 18-AUG-1981; GB-025150.
 PA (NATR) National Res Dev Corp.
 PA (WELL) Wellcome Foundation Ltd.
 PI Boothroyd JC, Cross GAM, Highfield PE, Winther MD, Rowlands DJ,
 PI Brown F, Harris TJR, Lowe PA;
 DR WPI; 82-26702E/14.
 DR N-PSDB; N20019.
 PT DNA corresp. to (part of) foot and mouth disease virus RNA - useful
 PT in prepn. of vaccines for producing antibodies against the virus
 PS Example; Fig 6; 57pp; English.
 CC The inventors claim a DNA molecule comprising a nucleotide sequence
 CC corresp. to all or a portion of foot-and-mouth disease virus RNA
 CC (FMDV). The DNA molecule is esp. for a precursor of FMDV capsid
 CC protein. It esp. codes for FMDV protein p88 and VP1-VP4. It may code
 CC for VP4, VP2, VP3 and VP1 contiguously. The inventors also claim a
 CC vaccine for stimulating prodn. of antibodies against FMDV in a
 CC mammal which comprises at least one of the above recombinant
 CC proteins produced by a host cell transformed with the DNA.
 SQ Sequence 78 AA;
 SQ 2 A; 2 R; 2 N; 0 D; 0 B; 3 C; 5 Q; 0 E; 0 Z; 4 G; 4 H;
 SQ 0 I; 8 L; 1 K; 2 M; 3 F; 8 P; 10 S; 15 T; 3 W; 2 Y; 4 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 ||||
 GALQTYWTWPKHVPPFFVSTMGNRTSLRGQTPVFWPSLMSPLPQNTCSTHTYQGLHSTTHSTLVLTCTSC
 10 X 20 X 30 40 50 60 70

15. US-08-249-182-7 (1-10)

R11715 HTLV-1 env. gp21 epitope encoded by ENV93 sequence

ID R11715 standard; Protein; 106 AA.
 AC R11715;
 DT 27-JUN-1991 (first entry)
 DE HTLV-1 env. gp21 epitope encoded by ENV93 sequence.
 KW Human T-cell leukaemia virus; HTLV-1; fusion protein; antibodies.
 OS Human T-cell leukaemia virus.
 PN EP-424748-A.
 PD 02-MAY-1991.
 PF 12-OCT-1990; 119560.
 PR 23-OCT-1989; US-425252.
 PA (HOFF) Hoffmann-La Roche AG.
 PI Buonagurio DA, Longiaru M;
 DR WPI; 91-126308/18.
 DR N-PSDB; Q11552.
 PT New nucleic acid encoding conserved epitope of HTLV gp.21 - and
 PT hybrids with other epitopes and the derived polypeptide(s) useful as
 PT immunoassay reagents for detecting specific HTLV antibodies in serum
 PS Claim 3; page 12; 31pp; English.
 CC This epitope, from the immunodominant conserved region of HTLV-1
 CC envelope (env) glycoprotein (gp)21, is encoded by the ENV93 gene
 CC construct. It constitutes residues 342-434 of the gp21 sequence.
 CC This ENV93 gene construct may be used alone or as a vehicle for the
 CC high level expression of other epitopes of HTLV-1 env as fusion
 CC proteins. In the 2nd case the ENV93 sequence is fused to a sequence

13. US-08-249-182-7 (1-10)
P20022 Sequence of a foot and mouth disease virus capsid

ID P20022 standard; Protein; 67 AA.
AC P20022;
DT 20-AUG-1992 (first entry)
DE Sequence of a foot and mouth disease virus capsid protein
DE encoded by a region of recombinant plasmid pFA61/t76
KW Vaccine; antibody; capsid protein; immunogen; antigen;
KW foot and mouth disease.
OS Foot and mouth disease virus.
PN EP--48455-A.
PD 31-MAR-1982.
PF 17-SEP-1981.
PR 18-SEP-1980; GB-030208.
PR 22-OCT-1980; GB-034130.
PR 27-NOV-1980; GB-038147.
PR 08-APR-1981; GB-011064.
PR 18-AUG-1981; GB-025150.
PA (NATR) National Res Dev Corp.
PA (WELL) Wellcome Foundation Ltd.
PI Boothroyd JC, Cross GAM, Highfield PE, Winther MD, Rowlands DJ,
PI Brown F, Harris TJR, Lowe PA;
DR WPI; 82-26702E/14.
DR N-PSDB; N20020.
PT DNA corresp. to (part of) foot and mouth disease virus RNA - useful
PT in prepn. of vaccines for producing antibodies against the virus
PS Example; Fig 7; 57pp; English.
CC The inventors claim a DNA molecule comprising a nucleotide sequence
CC corresp. to all or a portion of foot-and-mouth disease virus RNA
CC (FMDV). The DNA molecule is esp. for a precursor of FMDV capsid
CC protein. It esp. codes for FMDV protein p88 and VP1-VP4. It may code
CC for VP4, VP2, VP3 and VP1 contiguously. The inventors also claim a
CC vaccine for stimulating prodn. of antibodies against FMDV in a
CC mammal which comprises at least one of the above recombinant
CC proteins produced by a host cell transformed with the DNA.
SQ Sequence 67 AA;
SQ 1 A; 9 R; 1 N; 0 D; 0 B; 3 C; 3 Q; 3 E; 0 Z; 3 G; 3 H;
SQ 0 I; 3 L; 2 K; 0 M; 2 F; 8 P; 12 S; 7 T; 2 W; 0 Y; 5 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:15-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
Residue Identity = 40% Matches = 4 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```

      X      X
      VPPFENIELY
      ||    ||
TRSPVSFTCRPPPSVSELRFQWPGWSRSPRVSSGPHPKTLREQNSSKHVTLTTSCHSQERRVAGGT
      10      20      30      40      50      60

```

14. US-08-249-182-7 (1-10)
P20020 Sequence of a foot and mouth disease virus capsid

ID P20020 standard; Protein; 78 AA.
AC P20020;
DT 20-AUG-1992 (first entry)
DE Sequence of a foot and mouth disease virus capsid protein
DE encoded by a region of recombinant plasmid pFA61/t76
KW Vaccine; antibody; capsid protein; immunogen; antigen;
KW foot and mouth disease.
OS Foot and mouth disease virus.
PN EP--48455-A.
PD 31-MAR-1982.

CC expressed by a vector produced by a novel form of fusion PCR which
 CC enables fusion of heavy and light chains prior to vector ligation,
 CC avoiding the cumbersome separate cloning of fragments. This linker
 CC sequence moves the SpeI site, retains native IgG1 upper hinge region,
 CC retains original lamB sequence.
 SQ Sequence 28 AA;
 SQ 2 A; 0 R; 0 N; 2 D; 0 B; 1 C; 0 Q; 2 E; 0 Z; 0 G; 1 H;
 SQ 0 I; 2 L; 3 K; 0 M; 2 F; 5 P; 4 S; 3 T; 0 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:26-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 || ||
 EPKSCDKTHTSPAPAPPELLKSSFYFDY
 10 20

12. US-08-249-182-7 (1-10)
 R27562 Insert B to prevent steric hindrance & competition

ID R27562 standard; Protein; 34 AA.
 AC R27562;
 DT 26-FEB-1993 (first entry)
 DE Insert B to prevent steric hindrance & competition with E. coli coat.
 KW Dicistronic expression vector; fusion PCR; antibody; cDNA library;
 KW ss.
 OS Synthetic.
 PN W09215678-A.
 PD 17-SEP-1992.
 PF 27-FEB-1992; U01475.
 PR 01-MAR-1991; US-663442.
 PA (STRA-) STRATAGENE.
 PI Sorge JA;
 DR WPI; 92-331724/40.
 PT Prodn. of dicistronic DNA library used to make antibodies, etc. -
 PT includes forming 1st and 2nd PCR admixtures, subjecting them to
 PT PCR thermo-cycles, sepg. double stranded DNA, hybridising, etc.
 PS Disclosure; Page 100; 143pp; English.
 CC This peptide linker sequence is used to increase the distance of an
 CC expressed IgG polypeptide from the surface membrane of E. coli, which
 CC results in decreased steric hindrance and competition of the preselected
 CC polypeptide with the lipopolysaccharide coat of E. coli. The IgG is
 CC expressed by a vector produced by a novel form of fusion PCR which
 CC enables fusion of heavy and light chains prior to vector ligation,
 CC avoiding the cumbersome separate cloning of fragments. This linker
 CC sequence moves the SpeI site, retains original IgG1 upper hinge region,
 CC retains original lamB sequence.
 SQ Sequence 34 AA;
 SQ 2 A; 0 R; 0 N; 2 D; 0 B; 1 C; 0 Q; 2 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 2 L; 5 K; 0 M; 1 F; 6 P; 5 S; 5 T; 0 W; 1 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:26-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 || ||
 PKSCDKTHTPEKSTDKTHTSPAPAPPELLKSSFY
 10 20 30

FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH₂, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 24
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN W09318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT detection and vaccines against hepatitis C, HIV and HTLV
 PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are
 CC useful for the detection of antibodies to HCV, and/or HIV,
 CC and/or HTLV-I or II.
 SQ Sequence 24 AA;
 SQ 1 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 4 L; 0 K; 0 M; 0 F; 3 P; 7 S; 1 T; 0 W; 2 Y; 3 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:29-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      VPPFENIELY
      ||      ||
XVLYSPNVSVSPSSSTLLYPSLAX
      10      20
  
```

11. US-08-249-182-7 (1-10)

R27561 Insert A to prevent steric hindrance & competition

ID R27561 standard; Protein; 28 AA.
 AC R27561;
 DT 26-FEB-1993 (first entry)
 DE Insert A to prevent steric hindrance & competition with E. coli coat.
 KW Dicistronic expression vector; fusion PCR; antibody; cDNA library;
 KW ss.
 OS Synthetic.
 PN W09215678-A.
 PD 17-SEP-1992.
 PF 27-FEB-1992; U01475.
 PR 01-MAR-1991; US-663442.
 PA (STRA-) STRATAGENE.
 PI Sorge JA;
 DR WPI; 92-331724/40.
 PT Prodn. of dicistronic DNA library used to make antibodies, etc. -
 PT includes forming 1st and 2nd PCR admixtures, subjecting them to
 PT PCR thermo-cycles, sepg. double stranded DNA, hybridising, etc.
 PS Disclosure; Page 100; 143pp; English.
 CC This peptide linker sequence is used to increase the distance of an
 CC expressed IgG polypeptide from the surface membrane of E. coli, which
 CC results in decreased steric hindrance and competition of the preselected
 CC polypeptide with the lipopolysaccharide coat of E. coli. The IgG is

PN EP-507573-A.
 PD 07-OCT-1992.
 PF 02-APR-1992; 302882.
 PR 03-APR-1991; JP-071253.
 PA (SANW) SANWA KAGAKU KENKYUSHO CO.
 PI Kato B, Kurono M, Mitani T, Owaki H, Sato M, Sawai K;
 PI Takahashi H;
 DR WPI; 92-333790/41.
 PT Orally and nasally active motilin compsns. - for promoting
 PT motility of digestive tract
 PS Claim 2; Page 7; 8pp; English.
 CC The peptide is a motilin-like substance, pref. L-leucine-13-motilin-
 CC homoserine, which may be administered orally or nasally, as distinct
 CC from prior art compsns. which are only effective by injection.
 CC Cpds. contg. this peptide may be readily absorbed to promote
 CC motility of the digestive tract, and are used for treating
 CC dyskinesia or for causing contraction of the muscles after surgical
 CC operation. The compsn. may opt. contain a peptolytic enzyme
 CC inhibitor to prevent or reduce attack by trypsin or chymotrysin on
 CC the peptide.
 SQ Sequence 23 AA;
 SQ 0 A; 2 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 3 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 2 K; 0 M; 2 F; 0 P; 1 S; 1 T; 0 W; 1 Y; 1 V;
 SQ 6 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:25-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      VPPFENIELY
      | |  ||
      FVXIFTYXELRXQEKERXKGXS
      X      10      20
  
```

10. US-08-249-182-7 (1-10)

R41084 HTLV-I and HTLV-II peptide I-gp46-3.

ID R41084 standard; peptide; 24 AA.
 AC R41084;
 DT 22-MAR-1994 (first entry)
 DE HTLV-I and HTLV-II peptide I-gp46-3.
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;

DE IgG1 light chain fragment #3.
 KW Fd'; fragment; human; IgG1; variable region; Fab'; F(ab')2; T cell;
 KW gene module; binding activity; antibody.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Cross_links 5
 FT /note= "Cys involved in disulphide bridge with
 FT heavy chain"
 PN W09222324-A.
 PD 23-DEC-1992.
 PF 15-JUN-1992; U04976.
 PR 14-JUN-1991; US-714175.
 PA (XOMA) XOMA CORP.
 PI Better MD, Carroll S, Horwitz AH;
 DR WPI; 93-017909/02.
 DR N-PSDB; Q34567-72.
 PT Polynucleotide sequences encoding Fab' and F(ab')2 fragments -
 PT used to produce, e.g. antibody-ricin A chain immuno:toxin(s)
 PS Disclosure; Fig 1b; 92pp; English.
 CC The sequences given in R30876-78 represent Fd' fragments of human
 CC IgG1. These polypeptides represent a variable region, which when
 CC part of an Fab' or F(ab')2 are reactive with T cells. Six different
 CC Fd' gene modules were constructed. These gene modules were used to
 CC create Ig fragments which retain the full binding activity of the
 CC whole antibody.
 SQ Sequence 14 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 1 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 2 L; 0 K; 0 M; 2 F; 5 P; 1 S; 1 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:34-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      VPPFENIELY
      ||      ||
      TSPPCPAPELLFFP
      X      10
  
```

9. US-08-249-182-7 (1-10)

R27059 Motilin-like peptide.

ID R27059 standard; peptide; 23 AA.
 AC R27059;
 DT 03-MAR-1993 (first entry)
 DE Motilin-like peptide.
 KW Absorb; orally; nasally; motility; digestive tract; dyskinesia.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 3
 FT /label= Pro, Gly, Asn, Ser
 FT Misc_difference 8
 FT /label= Gly, Pro, Asn, Ser
 FT Misc_difference 11
 FT /label= Gln, Glu, Asp
 FT Misc_difference 13
 FT /note= "any amino acid other than Met"
 FT Misc_difference 19
 FT /label= Asn, Glu, Asp
 FT Misc_difference 22
 FT /label= Gln, Lys, Arg
 FT Modified_site 23
 FT /note= "homoSer, its lactone or any polypeptide

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 57% Matches = 4 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

      X      X 10
      VPPFENIELY
      ||  ||
PILQERPPLENR
      X 10 X

```

7. US-08-249-182-7 (1-10)

R30876 IgG1 light chain fragment #1.

ID R30876 standard; peptide; 14 AA.
 AC R30876;
 DT 10-MAY-1993 (first entry)
 DE IgG1 light chain fragment #1.
 KW Fd'; fragment; human; IgG1; variable region; Fab'; F(ab')2; T cell;
 KW gene module; binding activity; antibody.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Cross_links 2
 FT /note= "Cys involved in disulphide bridge with
 FT heavy chain"
 FT Cross_links 5
 FT /note= "Cys involved in disulphide bridge with
 FT heavy chain"
 PN W09222324-A.
 PD 23-DEC-1992.
 PF 15-JUN-1992; U04976.
 PR 14-JUN-1991; US-714175.
 PA (XOMA) XOMA CORP.
 PI Better MD, Carroll S, Horwitz AH;
 DR WPI; 93-017909/02.
 DR N-PSDB; 034567-72.
 PT Polynucleotide sequences encoding Fab' and F(ab')2 fragments -
 PT used to produce, e.g. antibody-ricin A chain immuno:toxin(s)
 PS Disclosure; Fig 1b; 92pp; English.
 CC The sequences given in R30876-78 represent Fd' fragments of human
 CC IgG1. These polypeptides represent a variable region, which when
 CC part of an Fab' or F(ab')2 are reactive with T cells. Six different
 CC Fd' gene modules were constructed. These gene modules were used to
 CC create Ig fragments which retain the full binding activity of the
 CC whole antibody.
 SQ Sequence 14 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 2 C; 0 Q; 1 E; 0 Z; 2 G; 0 H;
 SQ 0 I; 2 L; 0 K; 0 M; 0 F; 5 P; 0 S; 1 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:34-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 2.68
 Residue Identity = 40% Matches = 4 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      VPPFENIELY
      ||  ||
TCPPCPAPELLGGP
      X 10

```

8. US-08-249-182-7 (1-10)

R30878 IgG1 light chain fragment #3.

ID R30878 standard; peptide; 14 AA.
 AC R30878;
 DT 10-MAY-1993 (first entry)
 DE IgG1 light chain fragment #3.
 KW Fd'; fragment; human; IgG1; variable region; Fab'; F(ab')2; T cell;
 KW gene module; binding activity; antibody.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Cross_links 2
 FT /note= "Cys involved in disulphide bridge with
 FT heavy chain"
 FT Cross_links 5
 FT /note= "Cys involved in disulphide bridge with
 FT heavy chain"
 PN W09222324-A.
 PD 23-DEC-1992.
 PF 15-JUN-1992; U04976.
 PR 14-JUN-1991; US-714175.
 PA (XOMA) XOMA CORP.
 PI Better MD, Carroll S, Horwitz AH;
 DR WPI; 93-017909/02.
 DR N-PSDB; 034567-72.
 PT Polynucleotide sequences encoding Fab' and F(ab')2 fragments -
 PT used to produce, e.g. antibody-ricin A chain immuno:toxin(s)
 PS Disclosure; Fig 1b; 92pp; English.
 CC The sequences given in R30876-78 represent Fd' fragments of human
 CC IgG1. These polypeptides represent a variable region, which when
 CC part of an Fab' or F(ab')2 are reactive with T cells. Six different
 CC Fd' gene modules were constructed. These gene modules were used to
 CC create Ig fragments which retain the full binding activity of the
 CC whole antibody.
 SQ Sequence 14 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 2 C; 0 Q; 1 E; 0 Z; 2 G; 0 H;
 SQ 0 I; 2 L; 0 K; 0 M; 0 F; 5 P; 0 S; 1 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:34-PDT using FindSeq

IN 12 OCT 1988; LT 402587.
 PA (RENA/) Renard A J J.
 PI Renard AJJ, Thiry MEG;
 DR WPI; 90-140005/19.
 DR N-NSDB; Q04292.
 PT Immunogenic recombinant polypeptide -
 PT from fish haemorrhagic septicaemia virus for protective
 PT vaccines, and new DNA sequences, vectors, transformed hosts and
 PT antibodies
 PS Claim 1; Page 66; 99pp; French.
 CC Recombinant polypeptide sequence is included in all
 CC recombinant proteins covered by the invention.
 CC See also Q04308, Q04318-20, Q05332-42.
 SQ Sequence 455 AA;
 SQ 19 A; 19 R; 21 N; 29 D; 0 B; 12 C; 14 Q; 18 E; 0 Z; 25 G; 23 H;
 SQ 31 I; 31 L; 23 K; 9 M; 17 F; 27 P; 34 S; 45 T; 9 W; 17 Y; 32 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:56:58-PDT using FindSeq

Initial Score = 5 Optimized Score = 6 Significance = 3.57
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || ||| |
 MEWNTFFLVILIIIIKSTTPQITQRPVENISTYHADWDTPLYTHPSNCRDSSFVPIRPAQLRCPHEFEDIN
 10 20 X 30 X 40 50 60 70

 KGLVSVPTRIIH
 80

6. US-08-249-182-7 (1-10)

R34234 HTLV-I gp21 transmembrane glycoprotein fragment 9.

ID R34234 standard; peptide; 12 AA.
 AC R34234;
 DT 04-AUG-1993 (first entry)
 DE HTLV-I gp21 transmembrane glycoprotein fragment 9.
 KW Human T-cell leukaemia virus; hydrophilic; conjugate; aggregate;
 KW diagnosis; antibodies.
 OS Synthetic.
 PN W09306843-A.
 PD 15-APR-1993.
 PF 08-OCT-1992; U08405.
 PR 08-OCT-1991; US-771553.
 PA (UYDU-) UNIV DUKE.
 PI Haynes BF, Palker TJ;
 DR WPI; 93-134125/16.
 PT Antigenic determinant peptide(s) of HTLV envelope glyco:protein -
 PT useful for detecting anti-HTLV-I and -II antibodies and as
 PT vaccine against HTLV
 PS Claim 4; Page 11; 50pp; English.
 CC The sequence of peptide 9 corresponds to residues 411-422 from the
 CC HTLV-I gp21 transmembrane glycoprotein. When covalently linked
 CC to a carrier mol. the hydrophilic peptide can induce in a mammal the
 CC prodn. of high titres of antibodies to gp46 envelope glycoprotein from
 CC HTLV-I or -II. The peptide and carrier may be used in vaccines against
 CC HTLV-I or -II infection. The peptide may be used in a diagnostic
 CC assay to detect the presence and titre of anti-HTLV antibodies.
 CC See also R34225-57.
 SQ Sequence 12 AA;
 SQ 0 A; 2 R; 1 N; 0 D; 0 B; 0 C; 1 Q; 2 E; 0 Z; 0 G; 0 H;
 SQ 1 I; 2 L; 0 K; 0 M; 0 F; 3 P; 0 S; 0 T; 0 W; 0 Y; 0 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:00-PDT using FindSeq

Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
  || |  ||
FVPIFTYGELQRLQEKERNKGQ
  X      10      20

```

4. US-08-249-182-7 (1-10)

P80164 Biosynthetic multifunctional protein.

ID P80164 standard; protein; 200 AA.
 AC P80164;
 DT 16-NOV-1990 (first entry)
 DE Biosynthetic multifunctional protein.
 KW Biosynthetic multifunctional protein; biosynthetic antibody binding site;
 KW protein trailer; ricin.
 PN W08809344-A.
 PD 01-DEC-1988.
 PF 19-MAY-1988; U01737.
 PR 21-MAY-1987; US-052800.
 PA (CREA-) Creative Biomolecules Inc.
 PI Huston JS, Oppermann H;
 DR WPI; 88-353928/49.
 DR N-PSDB; N80190.
 PT Recombinant multifunctional protein - having antibody binding site and a
 PT sequence for biological activity, ion sequestering or binding to a
 PT solid support.
 PS Disclosure; 115pp; English.
 CC The sequence is a biosynthetic multifunctional protein including a
 CC biosynthetic antibody binding site and a ricin protein trailer linked
 CC via a spacer sequence.
 SQ Sequence 200 AA;
 SQ 18 A; 17 R; 14 N; 6 D; 0 B; 1 C; 9 Q; 13 E; 0 Z; 16 G; 4 H;
 SQ 16 I; 17 L; 1 K; 2 M; 11 F; 8 P; 9 S; 15 T; 0 W; 12 Y; 11 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:03-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.57
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                  ||||
VVG YRAGNSAYFFHPDNQEDAEATHLFTDVQNR YTFAGGNYDRLEQLAGNLRENIELGNGLPEEAISALY
  90      100      110      120      130 X 140 X 150

YYSTGGTQLPTLARSFIICIQMISEAARFQYIEGENRT
  160      170      180      190

```

5. US-08-249-182-7 (1-10)

R04576 Polypeptide recognised by neutralising anti-HSV an

ID R04576 standard; protein; 455 AA.
 AC R04576;
 DT 24-SEP-1990 (first entry)
 DE Polypeptide recognised by neutralising anti-HSV antibodies
 KW Fish haemorrhagic septicaemia virus (HSV);
 KW immunogenic recombinant polypeptide; ss
 OS synthetic.
 PN CA2000570-A.
 PD 12-APR-1990.
 PF 12-OCT-1989; .

FT /label= extracellular_domain
 FT /note= "soluble, immunogenic form of IFN-R"
 PN EP-563487-A.
 PD 06-OCT-1993.
 PF 31-MAR-1992; 400902.
 PR 31-MAR-1992; EP-400902.
 PA (EUBI-) LAB EURD BIOTECHNOLOGIE SA.
 PI Benoit P, Maguire D, Meyer F, Plavec I, Tovey MG;
 DR WPI; 93-312951/40.
 DR P-PSDB; R42635.
 PT Monoclonal antibody to human interferon type-I receptor - having
 PT neutralising activity against human type I interferon, used for
 PT therapy and diagnosis
 PS Disclosure; Fig 3; 21pp; English.
 CC Monoclonal antibodies produced against soluble forms of the human
 CC interferon alpha-beta receptor based on the full-length human IFN-R
 CC sequence are claimed. The antibodies are useful for treatment and
 CC prophylaxis of disorders involving cell proliferation and/or viral
 CC infection.
 SQ Sequence 557 AA;
 SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
 SQ 43 I; 44 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 42 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:33-PDT using FindSeq

Initial Score = 6 Optimized Score = 6 Significance = 4.46
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPETTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFLEKRNPGNHLKWKQIPDCENVKT
      260      270      280
  
```

3. US-08-249-182-7 (1-10)
 P82050 13-Leu motilin

ID P82050 standard; peptide; 22 AA.
 AC P82050;
 DT 19-OCT-1990 (first entry)
 DE 13-Leu motilin
 KW Motilin; monoclonal antibodies.
 OS synthetic.
 PN J63044896-A.
 PD 25-FEB-1988.
 PF 13-AUG-1986; 189811.
 PR 13-AUG-1986; JP-189811.
 PA (KYOW) Kyowa Hakko Kogyo KK.
 DR WPI; 88-04821/14.
 PT Monoclonal antibody recognising 13-leucine motilin - used to accelerate
 PT peristalsis in humans
 PS Claim 3; page 1; 5pp; Japanese.
 CC By replacing M13 with leucine the activity is increased and the motilin
 CC is useful in medicine. Monoclonal antibodies can be prepared from this
 CC motilin deriv. which are useful for detecting its presence.
 SQ Sequence 22 AA;
 SQ 0 A; 2 R; 1 N; 0 D; 0 B; 0 C; 3 Q; 3 E; 0 Z; 2 G; 0 H;
 SQ 1 I; 2 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:00-PDT using FindSeq

Initial Score = 5 Optimized Score = 5 Significance = 3.57

35. R43883	Human kappa immunoglobulin 1	443	4	4	2.68	0
36. A49833	autoantigen recognized by an	443	4	4	2.68	0
37. R43339	Completely humanised C4G1 Ig	449	4	4	2.68	0
38. R33311	Humanised MaE11 Version 1 (in	453	4	4	2.68	0
39. R30774	H52H4-160 murine anti-CD18 an	454	4	4	2.68	0
40. R42066	Human anti-HBs heavy chain.	459	4	4	2.68	0

1. US-08-249-182-7 (1-10)

R37449 Autotaxin peptide ATX 48.

ID R37449 standard; peptide; 10 AA.
AC R37449;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 48.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 48. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 10 AA;
SQ 0 A; 0 R; 2 N; 0 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 1 I; 1 L; 0 K; 0 M; 1 F; 2 P; 1 S; 0 T; 0 W; 1 Y; 0 V;
CC Retrieved by shears on Wed 21 Sep 94 11:58:47-PDT using FindSeq

Initial Score = 8 Optimized Score = 8 Significance = 6.25
Residue Identity = 80% Matches = 8 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

X      X
VPPFENIELY
||||| ||
SPPFENINLY
X      10

```

2. US-08-249-182-7 (1-10)

R42635 Human interferon receptor.

ID R42635 standard; Protein; 557 AA.
AC R42635;
DT 20-APR-1994 (first entry)
DE Human interferon receptor.
KW IFN-R; extracellular domain; monoclonal antibody; viral infection;
KW cell proliferation; allograft rejection; systemic lupus erythematosus;
KW psoriasis; multiple sclerosis; Behcet's Disease; aplastic anaemia;
KW immunodeficiency; measles virus; interferon-alpha-beta.
OS Homo sapiens.
FH Key Location/Qualifiers

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	1	1.12

Times:	CPU	Total Elapsed
	00:00:08.91	00:00:08.00

Number of residues:	482836
Number of sequences searched:	5543
Number of scores above cutoff:	2620

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 6 standard deviations above mean ****						
1. R37449	Autotaxin peptide ATX 48.	10	8	8	6.25	0
**** 4 standard deviations above mean ****						
2. R42635	Human interferon receptor.	557	6	6	4.46	0
**** 3 standard deviations above mean ****						
3. P82050	13-Leu motilin	22	5	5	3.57	0
4. P80164	Biosynthetic multifunctional	200	5	5	3.57	0
5. R04576	Polypeptide recognised by neu	455	5	6	3.57	0
**** 2 standard deviations above mean ****						
6. R34234	HTLV-I gp21 transmembrane gly	12	4	4	2.68	0
7. R30876	IgG1 light chain fragment #1.	14	4	4	2.68	0
8. R30878	IgG1 light chain fragment #3.	14	4	4	2.68	0
9. R27059	Motilin-like peptide.	23	4	4	2.68	0
10. R41084	HTLV-I and HTLV-II peptide I-	24	4	4	2.68	0
11. R27561	Insert A to prevent steric hi	28	4	4	2.68	0
12. R27562	Insert B to prevent steric hi	34	4	4	2.68	0
13. P20022	Sequence of a foot and mouth	67	4	4	2.68	0
14. P20020	Sequence of a foot and mouth	78	4	4	2.68	0
15. R11715	HTLV-1 env. gp21 epitope enco	106	4	5	2.68	0
16. A48491	twitching motility protein P	135	4	4	2.68	0
17. R11721	ENV93/HTLV-1-IIIB' fusion pro	158	4	5	2.68	0
18. R21796	Bet v I allergen of birch.	160	4	4	2.68	0
19. R21791	Aln g I allergen of alder.	160	4	5	2.68	0
20. R11717	ENV93/HTLV-1-II fusion protei	217	4	5	2.68	0
21. R37573	Partial human skeletal muscle	223	4	4	2.68	0
22. R11716	ENV93/HTLV-1-I fusion protein	245	4	5	2.68	0
23. R11720	ENV93/HTLV-1-IIIA fusion prot	249	4	5	2.68	0
24. R24186	Bovine RSV strain A 51908 M p	256	4	4	2.68	0
25. R13117	Cholera toxin A1 fragment.	258	4	4	2.68	0
26. R11719	ENV93/HTLV-1-III fusion prote	315	4	5	2.68	0
27. R33279	43 kD endoflagellum sheath pr	320	4	5	2.68	0
28. P93704	Sequence of the 65kd surface	325	4	4	2.68	0
29. R37572	Rabbit skeletal muscle ADP-ri	327	4	4	2.68	0
30. R20130	SEQ ID No. 5 of the constant	330	4	4	2.68	0
31. P81026	C region of H chain (gamma1)	330	4	4	2.68	0
32. R20129	SEQ ID No. 4 of the constant	337	4	4	2.68	0
33. R11718	ENV93/HTLV-1-II+I fusion prot	344	4	5	2.68	0
34. R11068	12D3 antigen sequence deduced	346	4	4	2.68	0

Results file us-08-249-182-7.res made by on Wed 21 Sep 94 12:11:22-PDT.

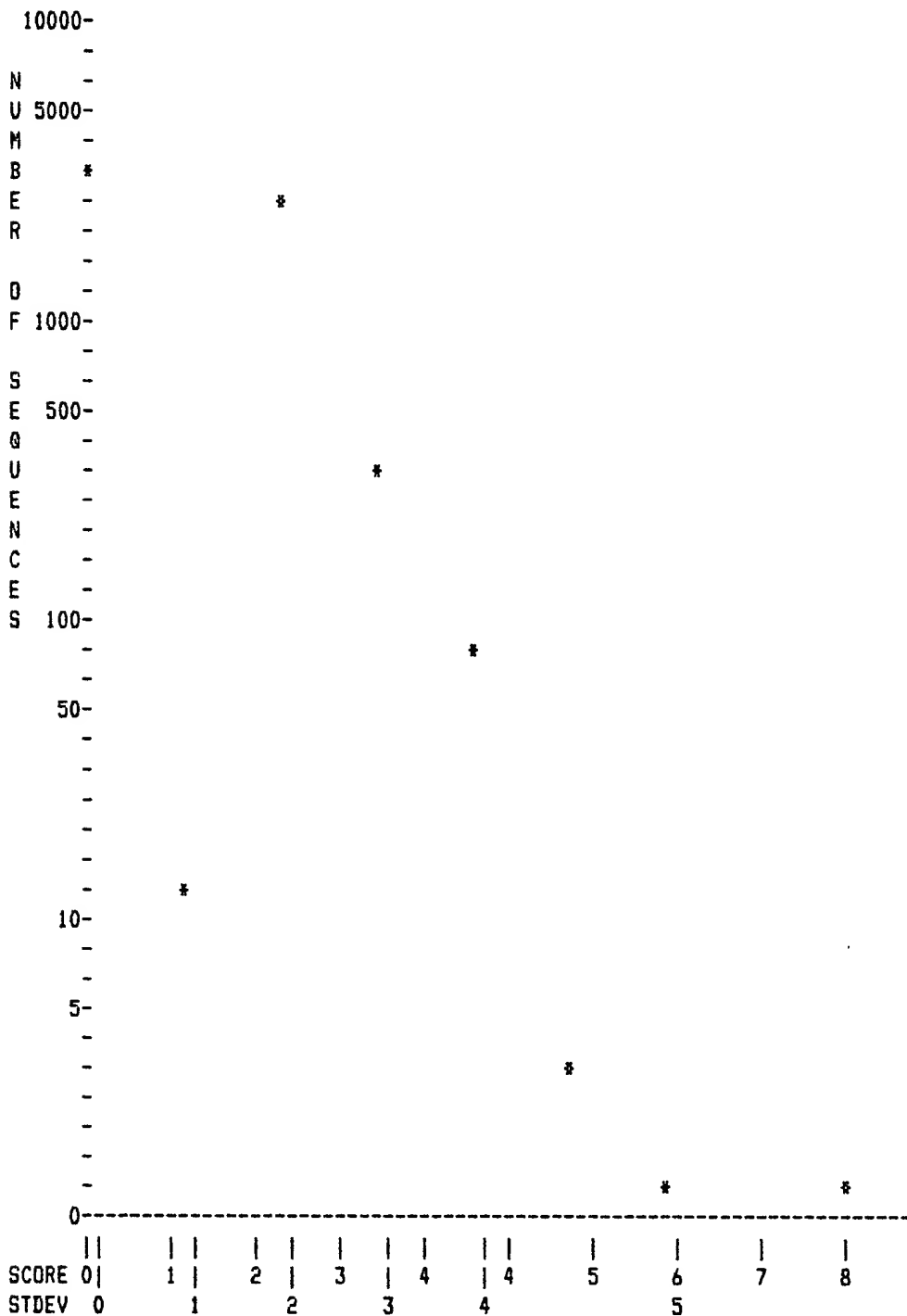
Query sequence being compared:US-08-249-182-7 (1-10)

Number of sequences searched: 5543

Number of scores above cutoff: 2620

Results of the initial comparison of US-08-249-182-7 (1-10) with:

File : /home/shears/loring/lorin*.pep



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

CC Fluids and for diagnosis of AIDS; ARC or pre-AIDS conditions
 CC It has application as a vaccine. See also P82715.
 SQ Sequence 34 AA;
 SQ 3 A; 2 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 1 E; 0 Z; 3 G; 0 H;
 SQ 3 I; 5 L; 2 K; 0 M; 0 F; 1 P; 1 S; 2 T; 2 W; 1 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:57:06-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 57% Matches = 4 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

          X      X
        YDVPWNETI
          ||||
RILAVERYLKDQQLLGIGCSGKLICTTAVPWNA
  10      20      30      X

```

15. US-08-249-182-6 (1-9)
 R44497 Sequence of the HIV-1 epitope gp41.

ID R44497 standard; Protein; 35 AA.
 AC R44497;
 DT 26-MAY-1994 (first entry)
 DE Sequence of the HIV-1 epitope gp41.
 KW Epitope; scFv; single chain antibody fragment; composite antibody;
 KW human immunodeficiency virus; HIV-1; surface protein; gp41.
 OS HIV-1.
 PN W09324630-A.
 PD 09-DEC-1993.
 PF 19-MAY-1993; AU0228.
 PR 22-MAY-1992; AU-002551.
 PA (AGEN-) AGEN LTD.
 PI Hillyard CJ, Hudson PJ, Lilley GG;
 DR WPI; 93-405821/50.
 PT Bifunctional recombinant protein - contains particle and analyte
 PT binding moieties, used in agglutination assays pref. on whole
 PT blood
 PS Example; Figure 7; 42pp; English.
 CC Epitopes of the surface protein gp41 from HIV1 and HIV2 virus types
 CC may be combined with epitopes from gp120 surface protein or p24
 CC core protein or substituted for the M2-FLAG epitope in scFv
 CC constructs or added to the scFv-M2 FLAG construct, thereby
 CC producing various bifunctional reagents capable of binding
 CC erythrocytes and serum antibodies which may be present in patient's
 CC serum. The sequences of M2-FLAG, HIV1 and HIV2 epitopes are given
 CC in R44496, R44497 and R44498 respectively.
 SQ Sequence 35 AA;
 SQ 4 A; 2 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 3 G; 0 H;
 SQ 3 I; 5 L; 2 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 1 Y; 2 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:45-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 50% Matches = 4 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

          X      X
        YDVPWNETI
          ||||
RILAVARYLKDQQLLGIGCSGKLICTTAVPWNAS
  10      20      30      X

```

> 0 <
 0| 0 IntelliGenetics
 > 0 <

SS.
 OS Synthetic.
 PN W09215678-A.
 PD 17-SEP-1992.
 PF 27-FEB-1992; U01475.
 PR 01-MAR-1991; US-663442.
 PA (STRA-) STRATAGENE.
 PI Sorge JA;
 DR WPI; 92-331724/40.
 PT Prodn. of dicistronic DNA library used to make antibodies, etc. -
 PT includes forming 1st and 2nd PCR admixtures, subjecting them to
 PT PCR thermo-cycles, sepg. double stranded DNA, hybridising, etc.
 PS Disclosure; Page 100; 143pp; English.
 CC This peptide linker sequence is used to increase the distance of an
 CC expressed IgG polypeptide from the surface membrane of E. coli, which
 CC results in decreased steric hindrance and competition of the preselected
 CC polypeptide with the lipopolysaccharide coat of E. coli. The IgG is
 CC expressed by a vector produced by a novel form of fusion PCR which
 CC enables fusion of heavy and light chains prior to vector ligation,
 CC avoiding the cumbersome separate cloning of fragments. This linker
 CC sequence moves the SpeI site, and retains the native IgG1 upper hinge
 CC region, retains original lamB sequence.
 SQ Sequence 29 AA;
 SQ 0 A; 0 R; 0 N; 4 D; 0 B; 1 C; 0 Q; 1 E; 0 Z; 1 G; 1 H;
 SQ 0 I; 0 L; 3 K; 0 M; 3 F; 2 P; 5 S; 3 T; 0 W; 4 Y; 1 V;
 CC Retrieved by shears on Wed 21 Sep 94 11:58:26-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 44% Matches = 4 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
      ||||
EPKSCDKTHTSYFYDVPDYGSKSSFYFDT
    10  X      20 X
  
```

14. US-08-249-182-6 (1-9)
 P82714 Peptide for detection of antibodies to HTLV-III.

ID P82714 standard; peptide; 34 AA.
 AC P82714;
 DT 23-NOV-1990 (first entry)
 DE Peptide for detection of antibodies to HTLV-III.
 KW HTLV-III; AIDS; ARC.
 OS synthetic.
 FH Key Location/Qualifiers
 FT Region 1..16
 FT /label=active fragment
 FT Region 17..35
 FT /label=active fragment
 PN ZA8704759-A.
 PD 05-JAN-1988.
 PF 01-JUL-1987; 004759.
 PR 11-SEP-1985; US-774644.
 PR 04-MAR-1986; US-837566.
 PR 02-APR-1986; US-847102.
 PR 10-FEB-1987; US-013014.
 PA (UNBI-) UTB Biomedical Inc.
 PI Wang CY, Wang JJG.;
 DR WPI; 88-147753/21.
 PT Peptide compsn. - used in detection of antibodies of HTLV-III.
 PS Claim 1; page 44; 48pp; English.
 CC This peptide, or its analogues or active fragments, is used in a
 CC compsn. for the detection of antibodies to HTLV-III in body

SS 2 Others;
CC Retrieved by shears on Wed 21 Sep 94 11:59:28-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
Residue Identity = 57% Matches = 4 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```
      X      X
      YDVPWNETI
      ||||
XDQQLLGIWGCGSGKHICTTNVPWNX
      10      20      X
```

12. US-08-249-182-6 (1-9)

P80566 Peptide region of human immunodeficiency virus-1 g

ID P80566 standard; protein; 27 AA.
AC P80566;
DT 07-NOV-1990 (first entry)
DE Peptide region of human immunodeficiency virus-1 gp41B1.
KW HIV-1; peptide; HIV-1 gp41B1; antibody; immunogen.
PN EP-284587-A.
PD 28-SEP-1988.
PF 28-MAR-1988; 850105.
PR 17-MAR-1987; SE-701294.
PR 18-MAY-1987; US-051726.
PA (VIR0-) Virovahl SA.
PI Vahine A, Svennerholm B, Rymo L, Jeansson S, Horal P;
DR WPI; 88-272997/39.
PT Synthetic peptide antigens for detection of HIV-1 infection -
PT also useful as immunogens in vaccine compsns.
PS Claim 5; page 24; 31pp; English.
CC The synthetic peptide correspond to regions of immunologically important
CC proteins of HIV-1. The peptide provides a superior, sensitive and
CC selective assay for the presence of antibodies to HIV-1. It may also be
CC used as an immunogen to elicit the prodn. of anti-HIV-1 antibodies.
CC X is the H of the NH2 group of the peptide, or an additional amino acid
CC selected to facilitate coupling of the peptide to a carrier protein.
CC Y is absent or Cys, and Z is OH or NH2.
SQ Sequence 27 AA;
SQ 2 A; 0 R; 4 N; 0 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 1 I; 1 L; 1 K; 2 M; 0 F; 1 P; 3 S; 2 T; 4 W; 0 Y; 1 V;
SQ 3 Others
CC Retrieved by shears on Wed 21 Sep 94 11:57:03-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
Residue Identity = 44% Matches = 4 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```
      X      X
      YDVPWNETI
      ||||
XTAVPWNASWSNKSLEQIWNMTWMYZ
      X      10      20
```

13. US-08-249-182-6 (1-9)

R27564 Insert D to prevent steric hindrance & competition

ID R27564 standard; Protein; 29 AA.
AC R27564;
DT 26-FEB-1993 (first entry)
DE Insert D to prevent steric hindrance & competition with E. coli coat.
KW Dicistronic expression vector; fusion PCR; antibody; cDNA library;

Initial Score = 4 Optimized Score = 4 Significance = 3.67
 Residue Identity = 44% Matches = 4 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
      ||||
XIWGCSGKLICTTAVPWNASX
    10 X    20

```

11. US-08-249-182-6 (1-9)

R41063 HIV-1 gp41 peptide (isolate ELI).

ID R41063 standard; peptide; 25 AA.
 AC R41063;
 DT 22-MAR-1994 (first entry)
 DE HIV-1 gp41 peptide (isolate ELI).
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;
 FT A (optional)= one or more amino acids, NH2 or
 FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH2, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 25
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN W09318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT detection and vaccines against hepatitis C, HIV and HTLV
 PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are
 CC useful for the detection of antibodies to HCV, and/or HIV,
 CC and/or HTLV-1 or II.
 SQ Sequence 25 AA;
 SQ 0 A; 0 R; 2 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 3 G; 1 H;

X X
 YDVPWNETI
 IIII
 IWGCSGKLICTTAVPWNAS
 10 X X

10. US-08-249-182-6 (1-9)

R41059 HIV-1 gp41 peptide (isolate HTLV-IIIB).

ID R41059 standard; peptide; 21 AA.
 AC R41059;
 DT 22-MAR-1994 (first entry)
 DE HIV-1 gp41 peptide (isolate HTLV-IIIB).
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;
 FT A (optional)= one or more amino acids, NH2 or
 FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH2, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 21
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN W09318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT detection and vaccines against hepatitis C, HIV and HTLV
 PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are
 CC useful for the detection of antibodies to HCV, and/or HIV,
 CC and/or HTLV-I or II.
 SQ Sequence 21 AA;
 SQ 2 A; 0 R; 1 N; 0 D; 0 B; 2 C; 0 Q; 0 E; 0 Z; 2 G; 0 H;
 SQ 2 I; 1 L; 1 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 0 Y; 1 V;
 SQ 2 Others;
 CC Retrieved by shears on Wed 21 Sep 94 11:59:28-PDT using FindSeq

DR WPI; 91-224505/31.
PT Peptide compsns. corresp. to envelope fragments of HTLV-1,2 - for
PT detecting antibodies to these viruses and diagnosing HIV and
PT adult T-cell leukaemia infections
PS Claim 19; Page 20; 27pp; English.
CC This peptide is one of 16 peptides useful for detecting antibodies to
CC HTLV or HIV viruses. The peptides correspond to partial sequences of
CC the HTLV virus designated gp21 and gp64, both part of gp61, which
CC defines the envelope protein of the HTLV-I or HTLV-II virus or
CC their analogues. The peptides can be amidated at the C-terminal.
CC This particular peptide is used for detecting antibodies to HIV-1 and
CC is included in a composition with at least one other peptide of the
CC invention which can detect HTLV-I and/or -II.
CC See R13184-R13193 and R13861-6.
SQ Sequence 19 AA;
SQ 2 A; 0 R; 1 N; 0 D; 0 B; 2 C; 0 Q; 0 E; 0 Z; 2 G; 0 H;
SQ 2 I; 1 L; 1 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:31-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
Residue Identity = 50% Matches = 4 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
        ||||
IWGCSGKLICTTAVPWNAS
    10 X      X

```

9. US-08-249-182-6 (1-9)

R06410 HTLV-1 corresponding peptide (VIII).

ID R06410 standard; protein; 19 AA.
AC R06410;
DT 21-DEC-1990 (first entry)
DE HTLV-1 corresponding peptide (VIII).
KW HTLV-1; HIV; antibodies; vaccines; polymers;
OS Synthetic.
PN WD9008162-A.
PD 26-JUL-1990.
PF 16-JAN-1990; U00260.
PR 13-JAN-1989; US-297635.
PA (UNBI-) UNITED BIOMED INC.
PI Yang CY;
DR WPI; 90-254015/33.
PT Synthetic peptide(s) corresponding to HTLV-1 and op. HIV - used
PT for detection of antibodies, in vaccines and for development of
PT antibodies
PS Claim 1 (VIII); Page 42; 52pp; English.
CC Peptides having specific immunoreactivity to antibodies to HTLV-1
CC and HIV comprise at least one peptide selected from those
CC represented in R06403-08; and this sequence on its own, or an analogue
CC of it in which amino acids may be added, deleted or substd, or segments,
CC mixts., conjugates or polymers of the peptides representes in R06409-12.
CC The peptides are safe, sensitive and specific in the detection of
CC antibodies. This peptide corresponds to sequences of HIV-1 or
CC HIV-2.
SQ Sequence 19 AA;
SQ 2 A; 0 R; 1 N; 0 D; 0 B; 2 C; 0 Q; 0 E; 0 Z; 2 G; 0 H;
SQ 2 I; 1 L; 1 K; 0 M; 0 F; 1 P; 2 S; 2 T; 2 W; 0 Y; 1 V;
CC Retrieved by shears on Wed 21 Sep 94 11:57:07-PDT using FindSeq

Initial Score = 4 Optimized Score = 4 Significance = 3.67
Residue Identity = 50% Matches = 4 Mismatches = 4

Loring
249182
Seq. IDs 1-11
(Pt. 2-2)

```
#####
#####
#####
#####
#####
#####
##### User: stic7!shears
#####
##### Title: u249_10a.res
#####
##### stic7
#####
##### Printed: Thu 11:22 Sep 22, 1994
#####
##### Job number: MT661-7-92
#####
#####
#####
#####
#####
> 0 <
0| |D IntelliGenetics
> 0 <
```

seq. 10
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_10a.res made by on Thu 22 Sep 94 10:28:32-PDT.

Query sequence being compared: US-08-249-182-10 (1-12)
Number of sequences searched: 42145
Number of scores above cutoff: 4700

Results of the initial comparison of US-08-249-182-10 (1-12) with:
Data bank : A-GeneSeq 15, all entries

100000-
-
N -
U50000-
M -
B -
E -

Scores:

Mean	Median	Standard Deviation
1	3	1.26

SEARCH STATISTICS

Parameter	Value
Similarity matrix	Unitary
Mismatch penalty	1
Gap penalty	1.00
Gap size penalty	0.05
Cutoff score	0
Randomization group	0
Initial scores to save	40
Optimized scores to save	0
Alignments to save	15
Display context	50
K-tuple	2
Joining penalty	20
Window size	5

PARAMETERS

[illegible]

00:00:25.95

00:00:26.00

Number of residues: 5287517
 Number of sequences searched: 42145
 Number of scores above cutoff: 4700

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37452	Autotaxin peptide ATX 102.	12	12	12	8.70	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
2. P93825	Neuron protein encoded by pCD	223	8	8	5.54	0
**** 3 standard deviations above mean ****						
3. R08086	Feline T-cell lymphotropic I	57	6	6	3.95	0
4. P82320	PAP-III isolated from biologi	145	6	6	3.95	0
5. R22403	Partial sequence of N-lipocor	304	6	6	3.95	0
6. R10689	Cephalosporin antibiotic bios	319	6	6	3.95	0
7. P91362	Human lipocortin-III.	323	6	6	3.95	0
8. P61523	Sequence of human lipocortin.	346	6	6	3.95	0
9. R06560	Human lipocortin obtained fro	346	6	6	3.95	0
10. P82063	Human lipocortin	346	6	6	3.95	0
11. P82062	Recombinant rat lipocortin	346	6	6	3.95	0
12. P82318	Lipocortin I isolated from bi	346	6	6	3.95	0
13. R22402	Human lipocortin.	363	6	6	3.95	0
14. P60657	Sequence of human lipocortin.	363	6	6	3.95	0
15. R41017	Insecticidal protein gene DRF	370	6	6	3.95	0
16. R14526	41.9 kD toxin.	370	6	6	3.95	0
17. P90400	Modified human lipocortin	387	6	6	3.95	0
18. R03725	Human placenta-derived coagul	672	6	6	3.95	0
19. R03726	Human placenta-derived coagul	786	6	6	3.95	0
20. R37309	Cardiac adenylyl cyclase.	1165	6	6	3.95	0
21. R32882	Cardiac adenylyl cyclase type	1184	6	6	3.95	0
22. P70647	Sequence of N-terminal apolip	2721	6	6	3.95	0
23. R23963	AFP-1 (Ala 2460 Val).	2783	6	6	3.95	0
24. R23962	AFP-1.	2783	6	6	3.95	0
25. R03180	Receptor fragment No376-390 c	15	5	6	3.16	0
26. R10873	Mammalian atrial natriuretic	17	5	5	3.16	0
27. R36481	DFII-4.5(36-60), a Dermatopha	25	5	5	3.16	0
28. R36482	DFII-4.3(45-70), a Dermatopha	26	5	5	3.16	0
29. R36483	DFII-15(51-77), a Dermatophag	27	5	5	3.16	0
30. R36490	DFII-22(36-63), a Dermatophag	28	5	5	3.16	0
31. R36489	DFII-21(33-60), a Dermatophag	28	5	5	3.16	0
32. P81534	Human insulin acceptor protei	29	5	6	3.16	0
33. P81533	Human insulin acceptor protei	37	5	6	3.16	0
34. R32204	Apple fruit PPD pSR8.	58	5	5	3.16	0
35. R37918	Cyn dI derived from clone 3 (138	5	5	3.16	0
36. R13904	Derf II allergen encoded by p	142	5	5	3.16	0
37. R13903	Derf II allergen encoded by p	142	5	5	3.16	0
38. R45140	Murine cytokine synthesis inh	157	5	5	3.16	0
39. R10159	pcD(SR alpha)-F115 mouse cyto	178	5	5	3.16	0

1. US-08-249-182-10 (1-12)

R37452 Autotaxin peptide ATX 102.

ID R37452 standard; peptide; 12 AA.
AC R37452;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 102.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 102. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 12 AA;
SQ 0 A; 1 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 1 H;
SQ 1 I; 2 L; 0 K; 0 M; 2 F; 0 P; 1 S; 1 T; 0 W; 0 Y; 0 V;

Initial Score = 12 Optimized Score = 12 Significance = 8.70
Residue Identity = 100% Matches = 12 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|||||
DIEHLTSLDFFR
X 10 X

2. US-08-249-182-10 (1-12)

P93825 Neuron protein encoded by pCDR31.

ID P93825 standard; protein; 223 AA.
AC P93825;
DT 12-JUN-1990 (first entry)
DE Neuron protein encoded by pCDR31.
KW Neuron polypeptide; pCDR31; paraneoplastic cerebellar degeneration.
FH Key Location/Qualifiers
FT Region 3..206
FT /note="region corresp. to tandem rpt region in N93188"
PN EP-297585-A.
PD 04-JAN-1989.
PF 30-JUN-1986; 110488.
PR 01-JUL-1987; US-068917.
PA (SLOK) Sloan-Kettering Inst.
PI Dropcho EJ, Chen YT, Posner JB, Old LJ;
DR WPI; 89-008858/02.
DR N-PSDB; N93188.

Initial Score = 8 Optimized Score = 8 Significance = 5.54
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

WWKT
220

R08086 Feline T-cell lymphotropic lentivirus of clone R5

```
Initial Score      =      6  Optimized Score =      6  Significance =  3.95
Residue Identity  =    50%  Matches          =      6  Mismatches   =      6
Gaps              =      0  Conservative Substitutions =      0
```

X 10 X
DIEHLTSLDFFR
|||| ||
ATNSQSCENVFSSXLPNGHNTVWEGVVFLPSIFWDLFXRVNRNCVLTSLFPFDLVIF
10 20 30 40 50 X

4. US-08-249-182-10 (1-12)

P82320 PAP-III isolated from biological fluid, used as an

ID P82320 standard; protein; 145 AA.
AC P82320;
DT 13-NOV-1990 (first entry)
DE PAP-III isolated from biological fluid, used as anticoagulant.
KW PAP-III; anticoagulant; anti-inflammatory agent; phospholipid;
KW phospholipase A2; disseminated intravascular coagulation;
KW deep vein thrombosis.
OS Homo sapiens.
PN W08805659-A.
PD 11-AUG-1988.
PF 05-FEB-1988; U00340.
PR 06-FEB-1987; US-011782.
PR 05-JUN-1987; US-059355.
PA (ZYMO-) Zymogenetics Inc; (UNIW) Univ of Washington.
PI Fujikawa K, Irani MH, Carter BLA;
DR WPI; 88-235049/33.
PT Human proteins having anticoagulant and antiinflammatory activity -
PT isolated from biological fluid by anion-exchange chromatographoc media.
PS Disclosure; p; English.
CC The biological fluid is a highly vacularised human tissue, e.g.
CC placenta. The product binds to phospholipid and inhibits phospholipase A2
CC The protein can substitute heparin or other anticoagulants in
CC the treatment of disseminated intravascular coagulation, deep vein
CC thrombosis, or other disorders. It also has antiinflammatory
CC properties.
SQ Sequence 145 AA;
SQ 10 A; 6 R; 2 N; 14 D; 0 B; 1 C; 4 Q; 9 E; 0 Z; 12 G; 1 H;
SQ 12 I; 18 L; 13 K; 4 M; 2 F; 0 P; 12 S; 9 T; 0 W; 9 Y; 6 V;
SQ 1 Others;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|| || ||

DDLKGDLSGHMVALVMKGAGTNEALIEILTTMKDIXQAYYTVYKKSLGDDISGETSGDFRKALLVLAVSRS
40 50 60 70 80 90 X 100

EIDLLDIRTEFKKRYGYSLSAISKSDTSGDYEITLLKICG
110 120 130 140

5. US-08-249-182-10 (1-12)

R22403 Partial sequence of N-lipocortin encoded by lambda

ID R22403 standard; Protein; 304 AA.
AC R22403;
DT 13-MAY-1992 (first entry)
DE Partial sequence of N-lipocortin encoded by lambda-NLip6-1X.
KW Antiinflammatory; arthritis; allergy; dermatology; placenta.
KW phospholipase A2 inhibitor.
OS Homo sapiens.
PN US5081019-A.
PD 14-JAN-1992.
PF 02-MAY-1990; 519256.
PR 10-JAN-1985; US-690146.
PR 15-MAR-1985; US-712376.
PR 14-AUG-1985; US-765877.
PR 05-SEP-1985; US-772892.

PR 08-MAR-1988; US-837019.
 PR 02-MAY-1990; US-519256.
 PA (BIOS) BIOGEN INC.
 PI Wallner BP, Pepinsky RB;
 DR WPI; 92-048295/06.
 DR N-PSDB; Q23218.
 PT Deoxyribonucleic acid encoding lipo corticoid polypeptide(s) -
 PT for producing antiinflammatory agents for treating arthritic,
 PT allergic, dermatologic ophthalmic and collagen diseases
 PS Disclosure; Fig 31; 56pp; English.
 CC The sequence was deduced from DNA obtd. from a plasmid isolated from
 CC a cDNA library prepd. from polyA+ mRNA from human macrophage cell
 CC line U937 (see Q23218 for details). Based on sequence homology
 CC with lipocortin it is estimated that ca. 200 bp are missing from the
 CC 5' end of the DNA sequence. Recombinant N-lipocortin prepd. using
 CC the DNA sequence has antiinflammatory activity and can be used for
 CC the treatment of arthritic, allergic, dermatologic, ophthalmic and
 CC collagen diseases.
 CC See also R22402.
 SQ Sequence 304 AA;
 SQ 19 A; 21 R; 9 N; 28 D; 0 B; 3 C; 13 Q; 20 E; 0 Z; 17 G; 2 H;
 SQ 22 I; 31 L; 30 K; 8 M; 6 F; 5 P; 21 S; 15 T; 1 W; 17 Y; 16 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||  || ||
QYDASELKASMKGLGTDEDSLIEIICSRTNQELQEINRVYKEMYKTDLEKDIISDTSGDFRKLHVALAKGRR
      80      90      100      110      120 X      130 X      140

AEDGSVIDYELIDQDARDLYDAGVKRKGTDPKWKISIMTE
      150      160      170      180
  
```

6. US-08-249-182-10 (1-12)

R10689 Cephalosporin antibiotic biosynthetic enzyme #3.

ID R10689 standard; Protein; 319 AA.
 AC R10689;
 DT 27-MAR-1991 (first entry)
 DE Cephalosporin antibiotic biosynthetic enzyme #3.
 KW cephalosporin; antibiotic;
 KW S-(L-alpha-aminoadipyl)-L-cysteiny-D-; valine synthetase;
 KW isopenicillin N synthetase; isopenicillin N epimerase;
 KW deacetoxycephalosporin C synthetase; beta-lactamase;
 KW deacetoxycephalosporin C hydroxylase.
 OS Lysobacter lactangenus.
 PN J02291274-A.
 PD 03-DEC-1990.
 PF 10-JAN-1990; 003762.
 PR 01-FEB-1989; JP-024710.
 PR 10-JAN-1990; JP-003762.
 PA (TAKE) TAKEDA CHEMICAL IND KK.
 DR WPI; 91-018854/03.
 DR N-PSDB; Q10190.
 PT Prepn. of cephalosporin series antibiotics - comprises culturing
 PT transformant of microbe transformed by plasmid contg. new DNA
 PT fragment
 PS Disclosure; Fig 15; 67pp; Japanese.
 CC This protein is encoded by DRF3 of the 23666bp sequence
 CC isolated from L.lactangenus and comprising the genes for the
 CC cephalosporin biosynthetic enzymes listed in the KEYWORDS. Plasmids

CC Containing at least one of Gln's 1-7 can be used to transform
 CC microbes, such as bacteria or yeast.
 CC See also Q10191-2.
 SQ Sequence 319 AA;
 SQ 27 A; 28 R; 9 N; 27 D; 0 B; 5 C; 10 Q; 19 E; 0 Z; 22 G; 4 H;
 SQ 12 I; 23 L; 7 K; 7 M; 22 F; 19 P; 26 S; 16 T; 2 W; 9 Y; 25 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 MTDSGIQIFDLDELEHGVRLDSEFRKSLFERGVFYVREDDSIKTEHAKAMDAVMDLFENGSAEQKNALRNLT
 10 X 20 X 30 40 50 60 70

 NV

7. US-08-249-182-10 (1-12)
 P91362 Human lipocortin-III.

ID P91362 standard; protein; 323 AA.
 AC P91362;
 DT 10-MAR-1993 (revised)
 DT 22-DEC-1989 (first entry)
 DE Human lipocortin-III.
 KW Human lipocortin-III; lambdaHLipo III-5; anti-inflammatory agent.
 OS Homo sapiens.
 PN EP-330396-A.
 PD 30-AUG-1989.
 PF 20-FEB-1989; 301603.
 PR 26-FEB-1988; US-160866.
 PA (BIDJ) Biogen, Inc.
 PI Wallner BP, Pepinsky RB, Browning JL.
 DR WPI; 89-250486/35.
 PT Human lipocortin cpds. III, IV, V, and VI - used in treatment of
 PT arthritic, allergic, dermatologic, ophthalmic and collagen disorders
 PT involving inflammatory processes.
 PS Claim 15; fig 3; 32pp; English.
 CC Human lipocortin-III was isolated from a lambda gt11 human lung cDNA
 CC library with rat lipocortin-III cDNA of lambda RLipo III-5 as probe.
 CC Lipocortins are anti-inflammatory agents and can be used to treat
 CC arthritic, allergic, dermatologic, ophthalmic, and collagen diseases.
 CC See also N90598, N90599, and P91363.
 SQ Sequence 323 AA;
 SQ 23 A; 19 R; 7 N; 30 D; 0 B; 3 C; 11 Q; 22 E; 0 Z; 22 G; 7 H;
 SQ 24 I; 36 L; 26 K; 6 M; 11 F; 6 P; 23 S; 21 T; 2 W; 12 Y; 12 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 VFDAKQLKKSMKGAGTNEDALIEILTTTRSRQMKDISQAYYTVYKKS LGDDITSETSGDFRKALLTLADGRR
 100 110 120 130 140 X 150 X 160

 DESLKLDEHLAKQDAQILYKAGENRWGTDDEKFTAILCLR
 170 180 190 200

8. US-08-249-182-10 (1-12)

ID P61523 standard; Protein; 346 AA.
 AC P61523;
 DT 08-AUG-1991 (first entry)
 DE Sequence of human lipocortin.
 KW Anti-inflammatory agent; steroid mediation; arthritis therapy;
 KW allergy.
 OS Homo sapiens.
 PN W08604094-A.
 PD 17-JUL-1986.
 PF 10-JAN-1986; U00027.
 PR 10-JAN-1985; US-690146.
 PR 15-MAR-1985; US-712376.
 PR 14-AUG-1985; US-765877.
 PR 05-SEP-1985; US-772892.
 PR 10-JAN-1986; US-929199.
 PA (BIOJ) BIOGEN NV.
 PI Wallner BP, Pepinsky BR, Garwin JL, Schindler DG, Huang KS;
 DR WPI; 86-196888/30.
 DR N-PSDB; N60556.
 PT New human lipocortin-like polypeptide(s) - are obtd. by
 PT recombinant DNA techniques and are antiinflammatory agents
 PT without usual side effects
 PS Claim 18; Page 79; 113pp; English.
 CC A human cDNA library of Escherichia coli cells contg. human
 CC macrophage cDNA sequences inserted into a phage cloning vector was
 CC screened using antisense oligonucleotide DNA probes corresp. to
 CC those regions of rat phospholipase A2 inhibitor protein having
 CC minimal nucleotide degeneracy (N60561-N60564) to obtain a sequence
 CC coding for human lipocortin (LC)-like polypeptide.
 SQ Sequence 346 AA;
 SQ 32 A; 19 R; 13 N; 26 D; 0 B; 4 C; 12 Q; 27 E; 0 Z; 20 G; 5 H;
 SQ 21 I; 34 L; 32 K; 9 M; 11 F; 8 P; 19 S; 22 T; 1 W; 11 Y; 20 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||  || ||
EFDADDELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDITS DTSGDFRNALLSLGKGDR
 120      130      140      150      160      170      180

SEDFGVNEDLADSDARALYEAGERRKGTDVNVFNILTTR
 190      200      210      220
  
```

9. US-08-249-182-10 (1-12)

R06560 Human lipocortin obtained from lambda LC.

ID R06560 standard; protein; 346 AA.
 AC R06560;
 DT 14-JAN-1991 (first entry)
 DE Human lipocortin obtained from lambda LC.
 KW Human lipocortin; lambda LC; inflammation reduction; arthritis;
 KW phospholipase A2.
 OS Homo sapiens.
 PN US4950646-A.
 PD 21-AUG-1990.
 PF 10-JAN-1986; 929199.
 PR 10-JAN-1985; US-690146.
 PR 15-MAR-1985; US-712376.
 PR 14-AUG-1985; US-765877.
 PR 05-SEP-1985; US-772892.

PA (BIOJ) BIOGEN NV.
 P1 Wallner BP, Pepinsky RB, Garwin JL, Schindler DG, Huang KS;
 DR WPI; 90-274549/36.
 DR N-PSDB; 005809.
 PT Pure fragment of human lipocortin - useful for reducing
 PT inflammation or for treating arthritis, etc.
 PS Disclosure; Fig 4; 51pp; English.
 CC cDNA can be operatively linked to expression control sequences and used
 CC in various mammalian or other eukaryotic or prokaryotic host cells to
 CC produce the human lipocortinlike polypeptide. The peptide was
 CC shown to inhibit exogenous phospholipase A2 in in vitro assays.
 CC The 37 kD protein can be used for reducing inflammation or treating
 CC arthritic, allergic, dermatologic, ophthalmic and collagen diseases
 CC and other diseases involving inflammation processes.
 CC See also 005805-25 and R07926-37.
 SQ Sequence 346 AA;
 SQ 34 A; 19 R; 13 N; 25 D; 0 B; 4 C; 13 Q; 26 E; 0 Z; 20 G; 6 H;
 SQ 21 I; 34 L; 32 K; 9 M; 11 F; 8 P; 18 S; 22 T; 1 W; 11 Y; 19 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 QFHADELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDTISDTSGDFRNALLSLAKGDR
 120 130 140 150 160 170 180
 SEDFGVNEDLADSDARALYEAGERRKGTDVNVFNTILTTR
 190 200 210 220

10. US-08-249-182-10 (1-12)

P82063 Human lipocortin

ID P82063 standard; protein; 346 AA.
 AC P82063;
 DT 22-OCT-1990 (first entry)
 DE Human lipocortin
 KW recombinant rat lipocortin; rat abdominal dropsy; human lipocortin ; ss.
 OS Homo sapiens.
 PN J63276497-A.
 PD 14-NOV-1988.
 PF 08-MAY-1987; JP-112145.
 PR 08-MAY-1987; J6-JP-112145.
 PA (YAMA) Yamanouchi Pharm KK.
 DR WPI; 88-365616/51.
 PT Recombinant rat lipocortin -
 PT obtd using gene derived from cells in rat abdominal dropsy
 PS Disclosure; ; Japanese.
 CC Human lipocortin has strong homology to rat lipocortin, differing
 CC at only 38 positions out of 346.(See P82063 for rat sequence).
 CC To obtain the rat sequence mRNA was isolated from rat abdominal
 CC dropsy cells and cDNA synthesised from it. Probes were synthesised
 CC according to the partial amino acid sequence of rat lipocortin
 CC (see N82038 and N82039). These were used to isolate plasmids contg
 CC the desired sequence of lipocortin DNA.
 SQ Sequence 346 AA;
 SQ 34 A; 19 R; 11 N; 26 D; 0 B; 5 C; 14 Q; 25 E; 0 Z; 20 G; 5 H;
 SQ 21 I; 34 L; 33 K; 9 M; 10 F; 8 P; 19 S; 21 T; 0 W; 13 Y; 19 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|| || ||

QFDADELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDTSDTSGDFRNALLSLAKGDR
120 130 140 150 160 170 180

SEDFGVNEDLADSDARALYEAGERRKGTDVNVFNTILTTR
190 200 210 220

11. US-08-249-182-10 (1-12)

P82062 Recombinant rat lipocortin

ID P82062 standard; protein; 346 AA.
AC P82062;
DT 22-OCT-1990 (first entry)
DE Recombinant rat lipocortin
KW recombinant rat lipocortin; rat abdominal dropsy; ss.
OS Rattus.
PN J63276497-A.
PD 14-NOV-1988.
PF 08-MAY-1987; JP-112145.
PR 08-MAY-1987; J6-JP-112145.
PA (YAMA) Yamanouchi Pharm KK.
DR WPI; 88-365616/51.
DR N-PSDB; N82025.
PT Recombinant rat lipocortin -
PT obtd using gene derived from cells in rat abdominal dropsy
PS Disclosure; ; Japanese.
CC mRNA was isolated from rat abdominal dropsy cells and cDNA
CC synthesised from it. Probes were synthesised according to the
CC partial amino acid sequence of rat lipocortin (see N82038 and
CC N82039). These were used to isolate plasmids contg the desired
CC sequence of lipocortin DNA.
CC See also P82063.
SQ Sequence 346 AA;
SQ 33 A; 19 R; 11 N; 24 D; 0 B; 6 C; 16 Q; 26 E; 0 Z; 19 G; 5 H;
SQ 21 I; 33 L; 31 K; 10 M; 9 F; 9 P; 16 S; 26 T; 0 W; 14 Y; 18 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|| || ||

QFDADELRAAMKGLGTDEDTLIEILTTRSNQQIREITRVYREELKRDIAKDTSDTSGDFRNALLALAKGDR
120 130 140 150 160 170 180

CEDMSVNQDLADTDARALYEAGERRKGTDVNVFNTILTTR
190 200 210 220

12. US-08-249-182-10 (1-12)

P82318 Lipocortin I isolated from biological fluid, used

ID P82318 standard; protein; 346 AA.
AC P82318;
DT 13-NOV-1990 (first entry)
DE Lipocortin I isolated from biological fluid, used as anticoagulant.
KW Lipocortin I; anticoagulant; anti-inflammatory agent; phospholipid;
KW phospholipase A2; disseminated intravascular coagulation;
KW deep vein thrombosis.
OS Homo sapiens.
PN W08805659-A.

PD 11-AUG-1988.
 PF 05-FEB-1988; U00340.
 PR 06-FEB-1987; US-011782.
 PR 05-JUN-1987; US-059355.
 PA (ZYMO-) Zymogenetics Inc; (UNIW) Univ of Washington.
 PI Fujikawa K, Irani MH, Carter BLA;
 DR WPI; 88-235049/33.
 PT Human proteins having anticoagulant and antiinflammatory activity -
 PT isolated from biological fluid by anion-exchange chromatographoc media.
 PS Disclosure; p; English.
 CC The biological fluid is a highly vacularised human tissue, e.g.
 CC placenta. The product binds to phospholipid and inhibits phopholipase A2.
 CC The protein can substitute heparin or other anticoagulants in
 CC the treatment of disseminated intravascular coagulation, deep vein
 CC thrombosis, or other disorders. It also has antiinflammatory
 CC properties.
 SQ Sequence 346 AA;
 SQ 33 A; 19 R; 14 N; 26 D; 0 B; 4 C; 13 Q; 26 E; 0 Z; 20 G; 5 H;
 SQ 21 I; 34 L; 32 K; 9 M; 11 F; 8 P; 19 S; 22 T; 1 W; 11 Y; 18 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||  || ||
QFDADELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDITSDTSGDFRNALLSLAKGDR
120      130      140      150      160      170      180

SEDFGVNEDLADSDARALYEAGERRKGTDVNVFNTILTTR
190      200      210      220

```

13. US-08-249-182-10 (1-12)

R22402 Human lipocortin.

ID R22402 standard; Protein; 363 AA.
 AC R22402;
 DT 13-MAY-1992 (first entry)
 DE Human lipocortin.
 KW Antiinflammatory; arthritis; allergy; dermatology; recombinant.
 KW Phospholipase A2 inhibitor.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label= signal_sequence
 FT Protein 20..363
 FT /label= lipocortin
 PN US5081019-A.
 PD 14-JAN-1992.
 PF 02-MAY-1990; 519256.
 PR 10-JAN-1985; US-690146.
 PR 15-MAR-1985; US-712376.
 PR 14-AUG-1985; US-765877.
 PR 05-SEP-1985; US-772892.
 PR 06-MAR-1986; US-837019.
 PR 22-FEB-1989; US-314316.
 PR 02-MAY-1990; US-519256.
 PA (BIOS) BIOGEN INC.
 PI Wallner BP, Pepinsky RB;
 DR WPI; 92-048295/06.
 DR N-PSDB; Q23216.
 PT Deoxyribonucleic acid encoding lipo corticoid polypeptide(s) -
 PT for producing antiinflammatory agents for treating arthritic,
 PT allergic, dermatologic ophthalmic and collagen diseases

PS Disclosure; Fig 4; 5pp; English.
CC The sequence deduced from DNA obtd. from 14 overlapping clones
CC isolated from a cDNA library prepd. from polyA+ mRNA from human
CC macrophage cell line U937 (see Q23216 for details). Recombinant
CC lipocortin prepd. using the DNA sequence has antiinflammatory
CC activity and can be used for the treatment of arthritic, allergic,
CC dermatologic, ophthalmic and collagen diseases.
CC See also R22403.

SQ Sequence 363 AA;
SQ 33 A; 20 R; 13 N; 28 D; 0 B; 4 C; 14 Q; 27 E; 0 Z; 21 G; 5 H;
SQ 21 I; 36 L; 34 K; 9 M; 14 F; 8 P; 22 S; 23 T; 1 W; 11 Y; 19 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||  || ||
QFDADELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDITSDTSGDFRNALLSLAKGDR
    140      150      160      170      180 X 190 X 200

SEDFGVNEDLADSDARALYEAGERRKGTQVNVFNTILTTR
    210      220      230      240
```

14. US-08-249-182-10 (1-12)

P60657 Sequence of human lipocortin.

ID P60657 standard; Protein; 363 AA.
AC P60657;
DT 08-AUG-1991 (first entry)
DE Sequence of human lipocortin.
KW Anti-inflammatory agent; steroid mediation; arthritis therapy;
KW allergy.
OS Homo sapiens.
PN W08604094-A.
PD 17-JUL-1986.
PF 10-JAN-1986; U00027.
PR 10-JAN-1985; US-690146.
PR 15-MAR-1985; US-712376.
PR 14-AUG-1985; US-765877.
PR 05-SEP-1985; US-772892.
PR 10-JAN-1986; US-929199.
PA (BIOJ) BIOGEN NV.
PI Wallner BP, Pepinsky BR, Garwin JL, Schindler DG, Huang KS;
DR WPI; 86-196888/30.
DR N-PSDB; N60555.
PT New human lipocortin-like polypeptide(s) - are obtd. by
PT recombinant DNA techniques and are antiinflammatory agents
PT without usual side effects
PS Claim 18; page 78-79 and Fig 4; 113pp; English.
CC A human cDNA library of Escherichia coli cells contg. human
CC macrophage cDNA sequences inserted into a phage cloning vector was
CC screened using antisense oligonucleotide DNA probes corresp. to
CC those regions of rat phospholipase A2 inhibitor protein having
CC minimal nucleotide degeneracy (N60561-N60564) to obtain a sequence
CC coding for human lipocortin (LC)-like polypeptide.
SQ Sequence 363 AA;
SQ 33 A; 20 R; 13 N; 28 D; 0 B; 4 C; 14 Q; 27 E; 0 Z; 21 G; 5 H;
SQ 21 I; 36 L; 34 K; 9 M; 14 F; 8 P; 22 S; 23 T; 1 W; 11 Y; 19 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 QFDADELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDIAKDI TSDTSGDFRNALLSLAKGDR
 140 150 160 170 180 X 190 X 200
 SEDFGVNEDLADSDARALVEAGERRKGTDVNVFNTILTTR
 210 220 230 240

15. US-08-249-182-10 (1-12)

R41017 Insecticidal protein gene ORF-2 prod.

ID R41017 standard; Protein; 370 AA.
 AC R41017;
 DT 25-MAR-1994 (first entry)
 DE Insecticidal protein gene ORF-2 prod.
 KW Caulobacter; plasmid; insecticidal protein; Bacillus thuringiensis;
 KW Bacillus sphaericus; larva; mosquito; Culex; Anopheles; Psorophoa;
 KW Mansonia; Aedes.
 OS Bacillus sphaericus strain 2297.
 PN J05211866-A.
 PD 24-AUG-1993.
 PF 05-JUN-1991; 160963.
 PR 06-JUN-1990; JP-148444.
 PA (SILM-) SILMARAN SD TANABAL.
 DR WPI; 93-298916/38.
 DR N-PSDB; 048714.
 PT Expression of insecticidal protein - by transforming Caulobacter
 PT with plasmid contg. gene coding for insecticidal protein
 PS Disclosure; Page 10-14; 27pp; Japanese.
 CC Caulobacter transformed with a plasmid contg. a gene encoding
 CC insecticidal protein derived from Bacillus thuringiensis or
 CC Bacillus sphaericus will proliferate in aq. environment.
 CC They may be consumed by larvae of mosquitoes and are lethal to
 CC Culex, Anopheles, Psorophoa, Mansonia and Aedes.
 SQ Sequence 370 AA;
 SQ 13 A; 14 R; 27 N; 19 D; 0 B; 3 C; 13 Q; 22 E; 0 Z; 22 G; 6 H;
 SQ 33 I; 22 L; 15 K; 8 M; 21 F; 22 P; 32 S; 40 T; 3 W; 21 Y; 14 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.95
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 SAPNGDINTEICSRNNGYFIFFPTDDGRVIANRHNGSVFTGEATSVVSDIYTCSP LQFFREFKRTMSTVY
 40 50 60 70 80 X 90 100

LAIQNPESATDVRALEPNSHELPSRLYFTNNIENNSNILI
 110 120 130 140

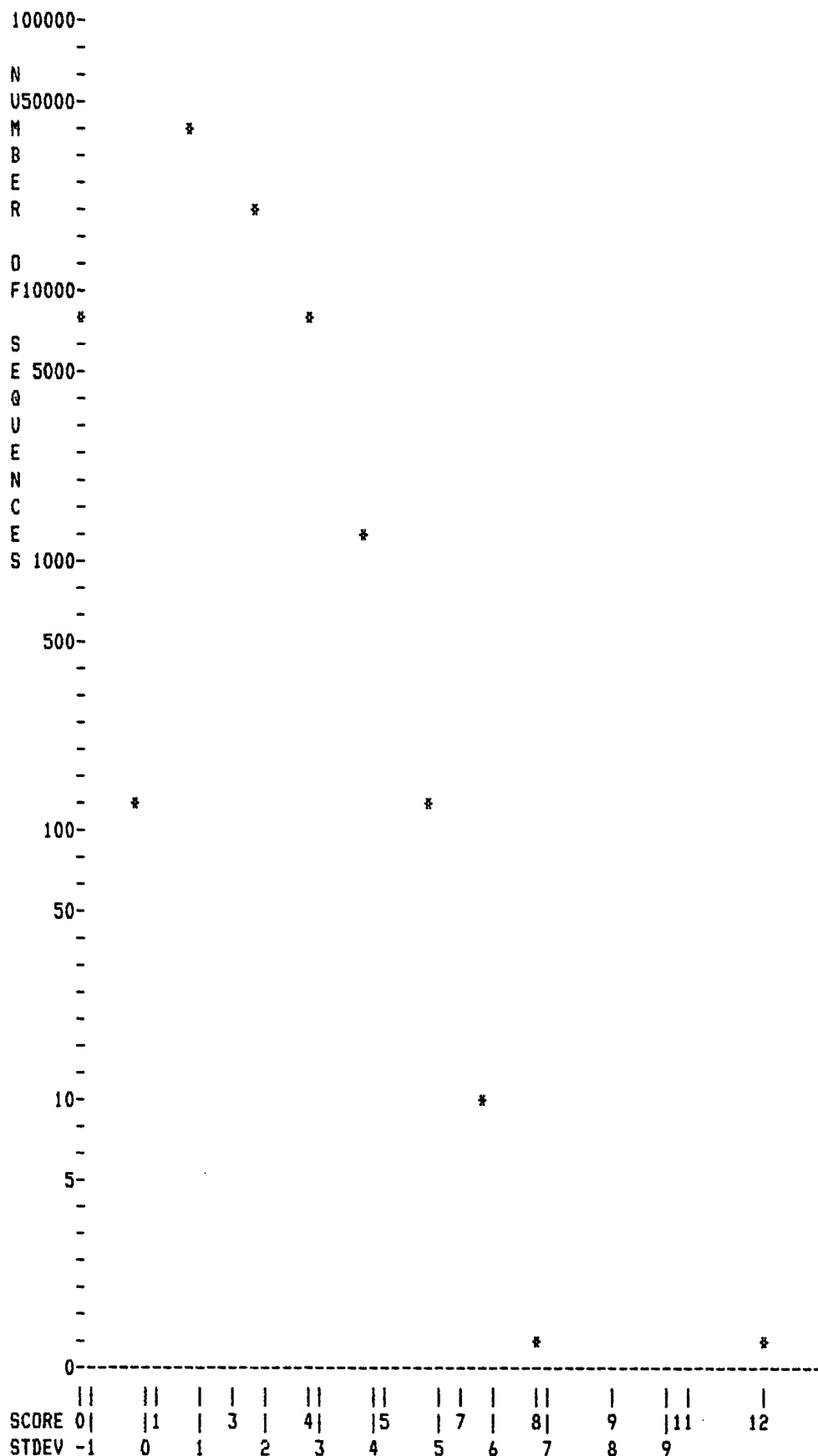
> 0 <
 0| 0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_10p.res made by on Thu 22 Sep 94 10:31:46-PDT.

Query sequence being compared:US-08-249-182-10 (1-12)
 Number of sequences searched: 70848
 Number of scores above cutoff: 4472

Results of the initial comparison of US 08 247-102-10 (1-12) with:
Data bank : PIR 41, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	1.03

Times:	CPU	Total Elapsed
	00:01:27.03	00:01:37.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	4472

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. A42329	autotaxin - human (fragments)	114	12	12	9.72	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
2. A29770	**** 5 standard deviations above mean **** cerebellar degeneration-relat	223	8	8	5.83	0
3. WMBEA1	**** 4 standard deviations above mean **** ribonucleoside-diphosphate re	321	7	7	4.86	0
4. S07743	cytochrome b - Paramecium tet	391	7	7	4.86	0
5. S18819	catalase (EC 1.11.1.6) - maiz	491	7	8	4.86	0
6. A36062	catalase (EC 1.11.1.6) - maiz	491	7	8	4.86	0
7. S10770	catalase (EC 1.11.1.6) - upla	492	7	8	4.86	0
8. S10395	catalase (EC 1.11.1.6) chain	492	7	8	4.86	0
9. S07124	catalase (EC 1.11.1.6) - swee	492	7	8	4.86	0
10. CSSY	catalase (EC 1.11.1.6) - soyb	492	7	8	4.86	0
11. CSPM	catalase (EC 1.11.1.6) - gard	494	7	8	4.86	0
12. S35908	ATP-dependent Clp proteinase	763	7	7	4.86	0
13. C20554	**** 3 standard deviations above mean **** hemocyanin LpIIa - Atlantic h	20	6	6	3.89	0
14. A20554	hemocyanin chain LpI - Atlant	24	6	6	3.89	0
15. F20554	hemocyanin LpIV - Atlantic ho	26	6	6	3.89	0
16. S31019	gene 74 protein - Mycobacteri	72	6	6	3.89	0
17. S21559	DNA gyrase chain B - Lyme dis	84	6	6	3.89	0
18. JN0145	hypothetical 13.6K protein (d	117	6	7	3.89	0
19. S33081	G4L protein - variola virus	124	6	6	3.89	0
20. A36844	H4L protein - variola virus (124	6	6	3.89	0
21. I42511	G4L protein - vaccinia virus	124	6	6	3.89	0
22. D35252	probable hemD protein - Bacil	130	6	6	3.89	0
23. S27507	hypothetical protein 1 - Baci	145	6	6	3.89	0
24. S41300	gene nimC protein - Bacteroid	163	6	6	3.89	0
25. S05361	hypothetical protein B (tdcR	184	6	6	3.89	0
26. S39221	hypothetical protein YBR0712	211	6	6	3.89	0

27. S27778	hypothetical protein - Caenor	234	6	6	3.89	0
28. S27789	hypothetical protein - Caenor	239	6	6	3.89	0
29. BVECA	appY protein - Escherichia co	243	6	6	3.89	0
30. BVECM5	M5 polypeptide - Escherichia	249	6	6	3.89	0
31. A45921	chorismate mutase (EC 5.4.99.	256	6	6	3.89	0
32. B42728	uroporphyrinogen-III synthase	262	6	6	3.89	0
33. A29633	arabinose operon regulatory p	281	6	6	3.89	0
34. RGEBA	arabinose operon regulatory p	281	6	6	3.89	0
35. JT0961	glutathione synthase (EC 6.3.	285	6	6	3.89	0
36. S32480	hypothetical protein B (L41 3	286	6	6	3.89	0
37. S13800	centrosomin A - mouse	289	6	6	3.89	0
38. S40580	arabinose operon regulatory p	292	6	6	3.89	0
39. RGECA	arabinose operon regulatory p	292	6	6	3.89	0
40. YXUNTP	thymidylate synthase (EC 2.1.	297	6	6	3.89	0

1. US-08-249-182-10 (1-12)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329
 ##status preliminary
 ##molecule_type protein
 ##residues 1-114 ##label STR
 ##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518; NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510; NCBIP:78509; NCBIP:78508; NCBIP:78503
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 114 #checksum 7335
 SEQUENCE

Initial Score = 12 Optimized Score = 12 Significance = 9.72
 Residue Identity = 100% Matches = 12 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      |||||
TEFLSNYLTVNDDITLVPGLGRDIEHLTSLDFFRVNSMOTVFVGYPGTFKGG@PLWITATKSPPFENINLY
      10      20 X 30 X 40      50      60      70

YDVPWNETIPEEV
      80

```

2. US-08-249-182-10 (1-12)

A29770 cerebellar degeneration-related protein - human

ENTRY A29770 #type complete
 TITLE cerebellar degeneration-related protein - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 03-Nov-1987 #sequence_revision 03-Nov-1987 #text_change 24-Jun-1993
 ACCESSIONS A29770

REFERENCE A27770
#authors Dropcho, E.J.; Chen, Y.T.; Posner, J.B.; Old, L.J.
#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:4552-4556
#title Cloning of a brain protein identified by autoantibodies from
a patient with paraneoplastic cerebellar degeneration.
#cross-references MUID:87260846
#accession A29770
##status preliminary
##molecule_type mRNA
##residues 1-223 ##label DR0
SUMMARY #length 223 #molecular-weight 27034 #checksum 9794
SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 5.83
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|| | ||||
LEDVDFQEDPNYPEDLDCWEDVDFLEDWRLLDMDFLEDVDLQEDIWLEDLDFRKMWIDWKTWI
150 160 170 180 190 200 210

WWKT
220

3. US-08-249-182-10 (1-12)

WMBEA1 ribonucleoside-diphosphate reductase (EC 1.17.4.1)

ENTRY WMBEA1 #type complete
TITLE ribonucleoside-diphosphate reductase (EC 1.17.4.1) small
chain - equine herpesvirus 1 (strain Ab4p)
ALTERNATE_NAMES ribonucleotide reductase small chain
ORGANISM #formal_name equine herpesvirus 1
#note host Equus caballus (domestic horse)
DATE 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change
04-Mar-1994
ACCESSIONS C36797
REFERENCE A36805
#authors Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
#submission submitted to GenBank, March 1992
#description The DNA sequence of equine herpesvirus-1.
#accession C36797
##molecule_type DNA
##residues 1-321 ##label TEL
##cross-references GB:M86664
REFERENCE A41831
#authors Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
#journal Virology (1992) 189:304-316
#title The DNA sequence of equine herpesvirus-1.
#cross-references MUID:92295566
#contents annotation; possible protein-coding frames
#note neither protein nor nucleotide sequence is given in this
paper
GENETICS
#gene 20
CLASSIFICATION #superfamily herpesvirus ribonucleoside-diphosphate reductase
small chain
KEYWORDS early protein; oxidoreductase
SUMMARY #length 321 #molecular-weight 36017 #checksum 5528
SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.86
Residue Identity = 58% Matches = 7 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

X      10 X
DIEHLTSLDFFR
||||| ||
MSIENSKEAALTAELSLAGAFFYTPECPDIEHLRSLSVANRWLDTDLPIISDDLKDVAKLTPAEREFYRFLFA
10      20      30      40      50      60      70

FLSAADDLVNVLNLGDLA
80      90

```

4. US-08-249-182-10 (1-12)

S07743 cytochrome b - *Paramecium tetraurelia* mitochondrio

```

ENTRY      S07743      #type complete
TITLE      cytochrome b - Paramecium tetraurelia mitochondrion (SGC6)
ORGANISM    #formal_name mitochondrion Paramecium tetraurelia
DATE        31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
            18-Jun-1993
ACCESSIONS  S07743
REFERENCE    S07725
#authors    Pritchard, A.E.; Seilhaner, J.J.; Mahalingam, R.; Sable,
            C.L.; Venuti, S.E.; Cummings, D.J.
#journal     Nucleic Acids Res. (1990) 18:173-180
#title       Nucleotide sequence of the mitochondrial genome of
            Paramecium.
#cross-references MUID:90174913
#accession    S07743
##molecule_type DNA
##residues     1-391 ##label PRI
##cross-references EMBL:X15917
##note         the translation of the nucleotide sequence is not given
            in this paper

GENETICS
#gene        cytB
#genome      mitochondrion
#genetic_code SGC6
#start_codon ATC
KEYWORDS     mitochondrion
SUMMARY      #length 391 #molecular-weight 46044 #checksum 3513
SEQUENCE

```

```

Initial Score      =      7  Optimized Score      =      7  Significance = 4.86
Residue Identity   =    58%  Matches                =      7  Mismatches   =    5
Gaps               =      0  Conservative Substitutions      =    0

```

```

X      10 X
DIEHLTSLDFFR
||| | |||
KNLRVSFHEVFSLFGFFTFMTIIVQLVSGTMLAFSSVPEPMLIPTVRDEEDIEDLYTDDFFWLHERGVDLIF
10      20      30      40      50      60      70      80

IFSYPHLLRKLYLVNVDLETEASWKSGVFSFLVFQVVVFF
90      100      110      120

```

5. US-08-249-182-10 (1-12)

S18819 catalase (EC 1.11.1.6) - maize

```

ENTRY      S18819      #type complete
TITLE      catalase (EC 1.11.1.6) - maize
ORGANISM    #formal_name Zea mays #common_name maize
DATE        22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
            22-Nov-1993
ACCESSIONS  S18819
REFERENCE    S18819

```

#journal Plant Physiol. (1991) 96:1379-1381
 #title Comparison of the cat2 complementary DNA sequences of a
 normal catalase activity line (W64A) and a high catalase
 activity line (R6-67) of maize.
 #accession S18819
 ##status preliminary
 ##residues 1-491 ##label GUA
 ##cross-references EMBL:X54819
 SUMMARY #length 491 #molecular-weight 56465 #checksum 8415
 SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
DSSLTVGARGPILLEDYHCEKLANFDRERIPERVVHARGASAKGFFEVDITHLTCADFLRAPGVQTPVIV
30      40      50      60      70      80      90      100

RFSTVIHERGSPETLRDPRGFAVKFYTREGNWDLVGNFP
      110      120      130      140
  
```

6. US-08-249-182-10 (1-12)
 A36062 catalase (EC 1.11.1.6) - maize

ENTRY A36062 #type complete
 TITLE catalase (EC 1.11.1.6) - maize
 ORGANISM #formal_name Zea mays #common_name maize
 DATE 16-Nov-1990 #sequence_revision 13-Jan-1993 #text_change
 23-Jun-1993
 ACCESSIONS A36062
 REFERENCE A36062
 #authors Bethards, L.A.; Skadsen, R.W.; Scandalios, J.G.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:6927
 #cross-references MUID:90370897
 #contents Erratum
 #accession A36062
 ##status preliminary
 ##molecule_type mRNA
 ##residues 1-491 ##label BET
 ##cross-references GB:J02976
 ##note the authors translated the codon AAC for residue 124 as
 Lys
 CLASSIFICATION #superfamily catalase
 KEYWORDS oxidoreductase
 SUMMARY #length 491 #molecular-weight 56506 #checksum 8423
 SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
DSSLTVGARGPILLEDYHCEKLANFDRERIPERVVHARGASAKGFFEVDITHLTCADFLRAPGVQTPVIV
30      40      50      60      70      80      90      100

RFSTVIHERGSPETLRDPRGFAVNFTREGNWDLVGNFP
      110      120      130      140
  
```

7. US-08-249-182-10 (1-12)

S10770 catalase (EC 1.11.1.6) - upland cotton

ENTRY S10770 #type complete
TITLE catalase (EC 1.11.1.6) - upland cotton
ORGANISM #formal_name Gossypium hirsutum #common_name upland cotton
DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change
21-Nov-1993
ACCESSIONS S10770
REFERENCE S10770
#authors Ni, W.; Turley, R.B.; Trelease, R.N.
#journal Biochim. Biophys. Acta (1990) 1049:219-222
#title Characterization of a cDNA encoding cottonseed catalase.
#cross-references MUID:90304227
#accession S10770
##status preliminary
##residues 1-492 ##label NIA
SUMMARY #length 492 #molecular-weight 56855 #checksum 5792
SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
SSLTVGPRGQYLLDYHLVEKLANFDRERIPERVVHARGASAKGFFEVDHSHLTCADFLRAPGVQTPVIV
30      40      50      60      70      80      90      100

RFSTVIHERGSPETLRDPRGFAVKFYTRECNFDLVGNFP
110     120     130     140
```

8. US-08-249-182-10 (1-12)

S10395 catalase (EC 1.11.1.6) chain 1 - upland cotton

ENTRY S10395 #type complete
TITLE catalase (EC 1.11.1.6) chain 1 - upland cotton
ORGANISM #formal_name Gossypium hirsutum #common_name upland cotton
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
28-Apr-1993
ACCESSIONS S10395
REFERENCE S10395
#authors Weiting, N.; Turley, R.B.; Trelease, R.N.
#submission submitted to the EMBL Data Library, March 1990
#accession S10395
##molecule_type mRNA
##residues 1-492 ##label WEI
##cross-references EMBL:X52135
CLASSIFICATION #superfamily catalase
KEYWORDS oxidoreductase
SUMMARY #length 492 #molecular-weight 56859 #checksum 5891
SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
SSLTVGPRGQYLLDYHLVEKLANFDRERIPERVVHARGASAKGFFDVTHDISHLTCADFLRAPGVQTPVIV
30      40      50      60      70      80      90      100

RFSTVIHERGSPETLRDPRGFAVKFYTRECNFDLVGNFP
```

9. US-08-249-182-10 (1-12)

S07124 catalase (EC 1.11.1.6) - sweet potato

ENTRY S07124 #type complete
 TITLE catalase (EC 1.11.1.6) - sweet potato
 ORGANISM #formal_name Ipomoea batatas #common_name sweet potato
 DATE 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change
 28-Apr-1993
 ACCESSIONS S07124
 REFERENCE S07124
 #authors Sakajo, S.; Nakamura, K.; Asahi, T.
 #journal Eur. J. Biochem. (1987) 165:437-442
 #title Molecular cloning and nucleotide sequence of full-length cDNA
 for sweet potato catalase mRNA.
 #cross-references MUID:87246622
 #accession S07124
 ##molecule_type mRNA
 ##residues 1-492 ##label SAK
 ##cross-references EMBL:X05549
 CLASSIFICATION #superfamily catalase
 KEYWORDS heme; heterotetramer; oxidoreductase
 SUMMARY #length 492 #molecular-weight 56985 #checksum 8615
 SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || ||| || |
 CALTVGSRGPILLEDYHVLVEKIQNFTIRERIPERVVHARGASAKGFFEVDITHLTCADFLRAPGVQTPILV
 30 40 50 60 70 80 90 100
 RFSTVIHERGSPETIRDPRGFAVKMYTRGGNWDLVGNFP
 110 120 130 140

10. US-08-249-182-10 (1-12)

CSSY catalase (EC 1.11.1.6) - soybean

ENTRY CSSY #type complete
 TITLE catalase (EC 1.11.1.6) - soybean
 ORGANISM #formal_name Glycine max #common_name soybean
 DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change
 30-Jun-1993
 ACCESSIONS S20999
 REFERENCE S20999
 #authors Allen, R.
 #submission submitted to the EMBL Data Library, May 1992
 #accession S20999
 ##molecule_type DNA
 ##residues 1-492 ##label ALL
 ##cross-references EMBL:Z12021
 GENETICS
 #introns 5/3; 38/1; 389/3; 419/3; 442/2; 473/3
 CLASSIFICATION #superfamily catalase
 KEYWORDS heme; oxidoreductase
 FEATURE
 65,104,138 #active_site His, Ser, Asn #status predicted\
 348 #binding_site heme iron (Tyr) (axial ligand) #status
 predicted
 SUMMARY #length 492 #molecular-weight 56847 #checksum 9929

SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
  SSLTVGSRGPILLEDYHLVEKLANFDRERIPERVVHARGASAKGFFEVDHSHLTCADFLRAPGVQTP LIV
  30          40          50          60          70          80          90         100

  RFSTVIHERGSPETLRDPRGFAVKFYTREGNFDLVGNFP
          110          120          130          140

```

11. US-08-249-182-10 (1-12)

CSPM catalase (EC 1.11.1.6) - garden pea

ENTRY CSPM #type complete
 TITLE catalase (EC 1.11.1.6) - garden pea
 ORGANISM #formal_name Pisum sativum #common_name garden pea
 DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 30-Jun-1993
 ACCESSIONS S18346; S15559
 REFERENCE S18346
 #authors Isin, S.H.; Allen, R.D.
 #journal Plant Mol. Biol. (1991) 17:1263-1265
 #title Isolation and characterization of a pea catalase cDNA.
 #cross-references MUID:92032793
 #accession S18346
 ##molecule_type mRNA
 ##residues 1-494 ##label ISI
 ##cross-references EMBL:X60169
 CLASSIFICATION #superfamily catalase
 KEYWORDS heme; oxidoreductase; peroxisome; tetramer
 FEATURE
 65,104,138 #active_site His, Ser, Asn #status predicted\
 348 #binding_site heme iron (Tyr) (axial ligand) #status predicted
 SUMMARY #length 494 #molecular-weight 57344 #checksum 6516
 SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.86
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
  SSLTVGSRGPILLEDYHLVEKLAQFDRERIPERVVHARGASAKGFFEVDHSHLTCADFLRAPGVQTPVIV
  30          40          50          60          70          80          90         100

  RFSTVIHERGSPETLRDPRGFAVKFYTREGNYDLVGNNFP
          110          120          130          140

```

12. US-08-249-182-10 (1-12)

S35908 ATP-dependent Clp proteinase (EC 3.4.21.-) chain c

ENTRY S35908 #type complete
 TITLE ATP-dependent Clp proteinase (EC 3.4.21.-) chain clpL - Lactococcus lactis subsp. lactis plasmid pUCL22
 ORGANISM #formal_name Lactococcus lactis subsp. lactis
 DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Feb-1994

ACCESSIONS S35908
 REFERENCE S35907
 #authors Huang, D.C.; Huang, X.F.; Novel, G.; Novel, M.
 #journal Mol. Microbiol. (1993) 7:957-965
 #title Two genes present on a transposon-like structure in
 Lactococcus lactis are involved in a Clp-family proteolytic
 activity.
 #accession S35908
 ##molecule_type DNA
 ##residues 1-763 ##label HUA
 ##cross-references EMBL:X62333

GENETICS
 #gene clpL
 #genome plasmid
 KEYWORDS ATP binding; hydrolase
 SUMMARY #length 763 #molecular-weight 86018 #checksum 3552
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.86
 Residue Identity = 58% Matches = 7 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||||| ||
KAADIKRVDRGTQKQIKKTHQKEKITATIDDVAGSVERLTGIPVSDMGANDIEHLKNLDKRLKVMVIGEDEA
      420      430      440      450      460      470 X      480

VKMVAKAIRNRAGFSEGDQPKGSFLFVGPTGVGKTELSQ
      490      500      510      520
  
```

13. US-08-249-182-10 (1-12)
 C20554 hemocyanin LpIIa - Atlantic horseshoe crab (fragme

ENTRY C20554 #type fragment
 TITLE hemocyanin LpIIa - Atlantic horseshoe crab (fragment)
 ORGANISM #formal_name Limulus polyphemus #common_name Atlantic
 horseshoe crab
 DATE 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change
 18-Jun-1993
 ACCESSIONS C20554
 REFERENCE A90478
 #authors Lamy, J.; Lamy, J.; Sizaret, P.Y.; Billiald, P.; Jolles, P.;
 Jolles, J.; Feldmann, R.J.; Bonaventura, J.
 #journal Biochemistry (1983) 22:5573-5583
 #title Quaternary structure of Limulus polyphemus hemocyanin.
 #accession C20554
 ##molecule_type protein
 ##residues 1-20 ##label LAM
 SUMMARY #length 20 #checksum 6229
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 3.89
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      DIEHLTSLDFFR
      |||||
TVKEKQSRLLPLFEHLTSLP
    10 X      20
  
```

14. US-08-249-182-10 (1-12)
 A20554 hemocyanin chain LpI - Atlantic horseshoe crab (fr

ENTRY A20554 #type fragment
 TITLE hemocyanin chain LpI - Atlantic horseshoe crab (fragment)
 ORGANISM #formal_name Linulus polyphemus #common_name Atlantic horseshoe crab
 DATE 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Jun-1993
 ACCESSIONS A20554
 REFERENCE A90478
 #authors Lamy, J.; Lamy, J.; Sizaret, P.Y.; Billiald, P.; Jolles, P.; Jolles, J.; Feldmann, R.J.; Bonaventura, J.
 #journal Biochemistry (1983) 22:5573-5583
 #title Quaternary structure of Linulus polyphemus hemocyanin.
 #accession A20554
 ##molecule_type protein
 ##residues 1-24 ##label LAM
 SUMMARY #length 24 #checksum 2682
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 3.89
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      |||||
TIKKKQASILALFEHLTSLPKQHI
    10 X      20 X

```

15. US-08-249-182-10 (1-12)

F20554 hemocyanin LpIV - Atlantic horseshoe crab (fragment)

ENTRY F20554 #type fragment
 TITLE hemocyanin LpIV - Atlantic horseshoe crab (fragment)
 ORGANISM #formal_name Linulus polyphemus #common_name Atlantic horseshoe crab
 DATE 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Jun-1993
 ACCESSIONS F20554
 REFERENCE A90478
 #authors Lamy, J.; Lamy, J.; Sizaret, P.Y.; Billiald, P.; Jolles, P.; Jolles, J.; Feldmann, R.J.; Bonaventura, J.
 #journal Biochemistry (1983) 22:5573-5583
 #title Quaternary structure of Linulus polyphemus hemocyanin.
 #accession F20554
 ##molecule_type protein
 ##residues 1-26 ##label LAM
 SUMMARY #length 26 #checksum 6863
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 3.89
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      |||||
TLKEKQDRILVLFHLTSLTKHQLPE
    10 X      20 X

```

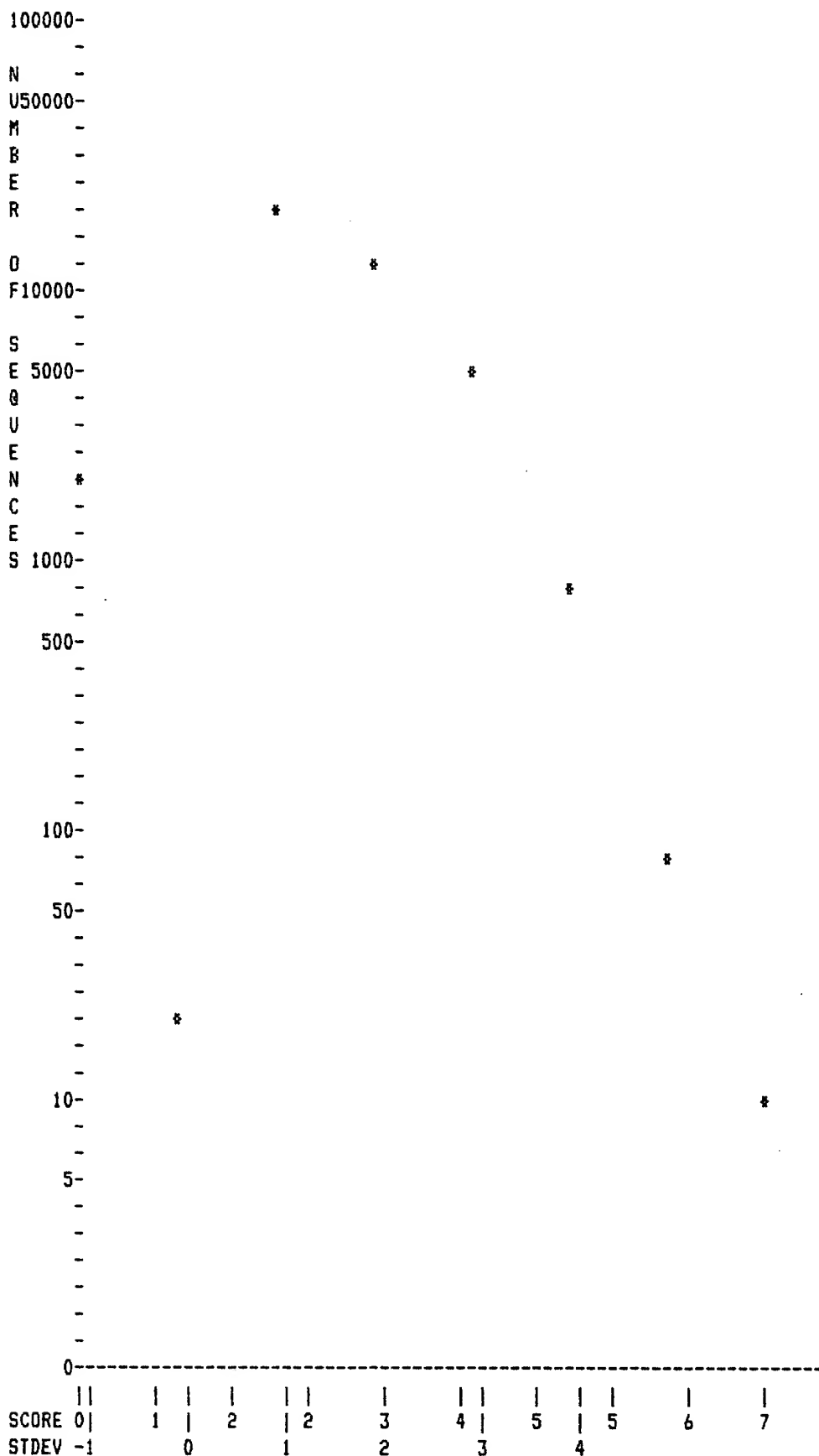
> 0 <
 0| |0 IntelliGenetics
 > 0 <

Query sequence being compared:US-08-249-182-10 (1-12)

Number of sequences searched: 36000

Number of scores above cutoff: 4217

Results of the initial comparison of US-08-249-182-10 (1-12) with:
Data bank : Swiss-Prot 28, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.94

Times:	CPU	Total Elapsed
	00:00:51.89	00:00:57.00

Number of residues:	12496420
Number of sequences searched:	36000
Number of scores above cutoff:	4217

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
1. RIR2_HSVEB	RIBONUCLEOSIDE-DIPHOSPHATE RE	321	7	7	5.31	0
2. CYB_PARTE	CYTOCHROME B (EC 1.10.2.2).	391	7	7	5.31	0
3. CAT2_MAIZE	CATALASE ISOZYME 2 (EC 1.11.1	491	7	8	5.31	0
4. CATA_SOYBN	CATALASE (EC 1.11.1.6).	492	7	8	5.31	0
5. CATA_IPOBA	CATALASE (EC 1.11.1.6).	492	7	8	5.31	0
6. CAT1_LYCES	CATALASE ISOZYME 1 (EC 1.11.1	492	7	8	5.31	0
7. CAT1_GOSHI	CATALASE ISOZYME 1 (EC 1.11.1	492	7	8	5.31	0
8. CATA_PEA	CATALASE (EC 1.11.1.6).	494	7	8	5.31	0
**** 4 standard deviations above mean ****						
9. VG74_BPMLS	GENE 74 PROTEIN (GP74).	72	6	6	4.25	0
10. YSCA_BACSU	HYPOTHETICAL 13.6 KD PROTEIN	117	6	7	4.25	0
11. VG04_VARV	PROTEIN G4.	124	6	6	4.25	0
12. VG04_VACCC	PROTEIN G4.	124	6	6	4.25	0
13. YME1_BACSU	HYPOTHETICAL 17.2 KD PROTEIN	145	6	6	4.25	0
14. YKA8_CAEEL	HYPOTHETICAL 27.3 KD PROTEIN	239	6	6	4.25	0
15. APPY_ECOLI	APPY PROTEIN (M5 POLYPEPTIDE)	243	6	6	4.25	0
16. CHMU_YEAST	CHORISMATE MUTASE (EC 5.4.99.	256	6	6	4.25	0
17. HEM4_BACSU	PUTATIVE UROPORPHYRINOGEN-III	262	6	6	4.25	0
18. ARAC_SALTY	ARABINOSE OPERON REGULATORY P	281	6	6	4.25	0
19. ARAC_CITFR	ARABINOSE OPERON REGULATORY P	281	6	6	4.25	0
20. YL44_KLULA	HYPOTHETICAL 31.8 KD PROTEIN	286	6	6	4.25	0
21. CENA_MOUSE	CENTROSOMIN A.	289	6	6	4.25	0
22. ARAC_ECOLI	ARABINOSE OPERON REGULATORY P	292	6	6	4.25	0
23. TYSY_PNECA	THYMIDYLATE SYNTHASE (EC 2.1.	297	6	6	4.25	0

25.	DIF_BOVIN	OSTEOINDUCTIVE FACTOR PRECURS	299	6	6	4.25	0
26.	ANX3_HUMAN	ANNEXIN III (LIPOCORTIN III)	323	6	6	4.25	0
27.	ANX3_RAT	ANNEXIN III (LIPOCORTIN III)	324	6	6	4.25	0
28.	VS09_ROTTEL	GLYCOPROTEIN VP7 (SEROTYPE-SP	326	6	6	4.25	0
29.	VG16_BPP7A	ENCAPSIDATION PROTEIN (LATE P	332	6	6	4.25	0
30.	VG16_BPPH2	ENCAPSIDATION PROTEIN (LATE P	332	6	6	4.25	0
31.	RIR2_HSV23	RIBONUCLEOSIDE-DIPHOSPHATE RE	337	6	6	4.25	0
32.	ANX2_MOUSE	ANNEXIN II (LIPOCORTIN II) (C	338	6	6	4.25	0
33.	ANX2_HUMAN	ANNEXIN II (LIPOCORTIN II) (C	338	6	6	4.25	0
34.	ANX2_CHICK	ANNEXIN II (LIPOCORTIN II) (C	338	6	6	4.25	0
35.	ANX2_BOVIN	ANNEXIN II (LIPOCORTIN II) (C	338	6	6	4.25	0
36.	RIR2_HSV1K	RIBONUCLEOSIDE-DIPHOSPHATE RE	339	6	6	4.25	0
37.	ANXB_XENLA	ANNEXIN II TYPE I (LIPOCORTIN	339	6	6	4.25	0
38.	ANX2_XENLA	ANNEXIN II TYPE II (LIPOCORTI	339	6	6	4.25	0
39.	RIR2_HSV11	RIBONUCLEOSIDE-DIPHOSPHATE RE	340	6	6	4.25	0
40.	YEJE_ECOLI	HYPOTHETICAL 38.1 KD PROTEIN	341	6	7	4.25	0

1. U5-08-249-182-10 (1-12)

RIR2_HSVB RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE SMALL CHAIN (

ID RIR2_HSVB STANDARD; PRT: 321 AA.
AC P28847;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE SMALL CHAIN (EC 1.17.4.1)
DE (RIBONUCLEOTIDE REDUCTASE).
GN 20.
OS EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 92295566
RA TELFORD E.A.R., WATSON M.S., MCBRIDE K., DAVISON A.J.;
RL VIROLOGY 189:304-316(1992).
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: 2'DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
CC THIOREDOXIN.
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -!- COFACTOR: CONTAINS TWO IRON IONS.
CC -!- SIMILARITY: HIGH TO OTHER EUKARYOTIC, PROKARYOTIC, AND VIRAL
CC RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE SMALL CHAINS.
DR EMBL; M86664; HENSECOMG.
DR PIR; C36797; WMBEA1.
DR PROSITE; PS00368; RIBORED_SMALL.
KW OXIDOREDUCTASE; DNA REPLICATION; IRON; EARLY PROTEIN.
FT METAL 78 78 IRON 1 (BY SIMILARITY).
FT METAL 108 108 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 171 171 IRON 2 (BY SIMILARITY).
FT METAL 205 205 IRON 2 (BY SIMILARITY).
FT METAL 208 208 IRON 2 (BY SIMILARITY).
FT ACT_SITE 115 115 BY SIMILARITY.
SQ SEQUENCE 321 AA; 36017 MW; 522774 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.31
Residue Identity = 58% Matches = 7 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
|||||

MSIENSREAAALTAELSLAGAFFYITECPDIEHLRSLSVANKWLDLCPISDULKDVAKLTPAEREFTYRPLFA
10 20 30 40 50 60 70

FLSAADDLVNINLGLSLA
80 90

2. US-08-249-182-10 (1-12)
CYB_PARTE CYTOCHROME B (EC 1.10.2.2).

ID CYB_PARTE STANDARD; PRT; 391 AA.
AC P15585;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CYTOCHROME B (EC 1.10.2.2).
GN COB OR CYTB.
OS PARAMECIUM TETRAURELIA.
OG MITOCHONDRION.
OC EUKARYOTA; PROTOZOA; CILIOPHORA; CILIATA; HOLOTRICHA; HYMENOSTOMATIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=STOCK 51;
RM 90174913
RA PRITCHARD A.E., SEILHAMER J.J., MAHALINGAM R., SABLE C.L.,
RA VENUTI S.E., CUMMINGS D.J.;
RL NUCLEIC ACIDS RES. 18:173-180(1990).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: $QH(2) + 2 \text{ FERRICYTOCHROME C} = Q +$
CC $2 \text{ FERROCYTOCHROME C.}$
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
DR EMBL; X15917; MIPAGEN.
DR PIR; S07743; S07743.
DR PROSITE; PS00192; CYTOCHROME_B_HEME.
DR PROSITE; PS00193; CYTOCHROME_B_Q0.
KW ELECTRON TRANSPORT; MITOCHONDRION; RESPIRATORY CHAIN; TRANSMEMBRANE;
KW HEME.
FT METAL 72 72 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 86 86 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 173 173 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 187 187 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 391 AA; 46044 MW; 895845 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.31
Residue Identity = 58% Matches = 7 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10 X
DIEHLTSLDFFR
||| | |||
KNLRVSFHEVFSLFGFFTFMTIIVQLVSGTMLAFSSVPEPHLIPTVRDEEDIEDLYTDDFFWLHERGVDLIF
10 20 30 40 50 60 70 80

IFS VFHLLRKLYLNVFDLETEASWKSGVFSFLVFQVVVFF
90 100 110 120

3. US-08-249-182-10 (1-12)
CAT2_MAIZE CATALASE ISOZYME 2 (EC 1.11.1.6).

ID CATA_MAIZE STANDARD; PRT; 491 AA.
 AC P12365;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE CATALASE ISOZYME 2 (EC 1.11.1.6).
 GN CAT2.
 OS ZEA MAYS (MAIZE).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
 OC CYPERALES; GRAMINEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W64A;
 RA GUAN L., RUZSA S., SKADSEN R.W., SCANDALIOS J.G.;
 RL PLANT PHYSIOL. 96:1379-1381(1991).
 RN [2]
 RP PRELIMINARY SEQUENCE FROM N.A.
 RC STRAIN=R6-67;
 RM 88016183
 RA BETHARDS L.A., SKADSEN R.W., SCANDALIOS J.G.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 84:6830-6834(1987).
 RN [3]
 RP REVISIONS TO C-TERMINAL SEQUENCE.
 RM 90370897
 RA BETHARDS L.A., SKADSEN R.W., SCANDALIOS J.G.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:6927-6927(1990).
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
 CC PEROXIDE AND OTHER ACTIVE OXYGEN SPECIES.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: HEME GROUP.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL OR CYTOPLASMIC.
 CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
 DR EMBL; X54819; ZMCAT2R.
 DR EMBL; J02976; ZMCAT2.
 DR PIR; A36062; A36062.
 DR PIR; S18819; S18819.
 DR PROSITE; PS00437; CATALASE_1.
 DR PROSITE; PS00438; CATALASE_2.
 KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
 KW PEROXISOME; MULTIGENE FAMILY.
 FT ACT_SITE 64 64 BY SIMILARITY.
 FT ACT_SITE 137 137 BY SIMILARITY.
 FT BINDING 347 347 PROXIMAL HEME LIGAND (BY SIMILARITY).
 FT VARIANT 296 296 T -> R.
 SQ SEQUENCE 491 AA; 56465 MW; 1188896 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || ||| || |
 DSSLTVGARGPILLEDYHCEKLANFDRERIPERVVHARGASAKGFFEVDITHLTCADFLRAPGVQTPVIV
 30 40 50 60 70 80 90 100
 RFSTVIHERGSPETLRDPRGFAVKFYTREGNWDLVGNNFP
 110 120 130 140

4. US-08-249-182-10 (1-12)
 CATA_SOYBN CATALASE (EC 1.11.1.6).

ID CATA_SOYBN STANDARD; PRT; 492 AA.

DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
 DE CATALASE (EC 1.11.1.6).
 OS GLYCINE MAX (SOYBEAN).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
 OC FABACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. D&PL 415;
 RA ISIN S.H., ALLEN R.D.;
 RL PLANT MOL. BIOL. 0:0-0(1992).
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
 CC PEROXIDE.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: HEME GROUP.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
 CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
 DR EMBL; Z12021; GMCATLG.
 DR PIR; S20999; CSSY.
 DR PROSITE; PS00437; CATALASE_1.
 DR PROSITE; PS00438; CATALASE_2.
 KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
 KW PEROXISOME.
 FT ACT_SITE 65 65 BY SIMILARITY.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 492 AA; 56847 MW; 1224278 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || ||| || |
 SSSLTVGSRGPILLEDYHLVEKLANFDRERIPERVVHARGASAKGFFEVDHISLTCADFLRAPGVQTPPLIV
 30 40 50 60 70 80 90 100
 RFSTVIHERGSPETLRDPRGFAVKFYTREGNFDLVGNFP
 110 120 130 140

5. US-08-249-182-10 (1-12)

CATA_IPOBA CATALASE (EC 1.11.1.6).

ID CATA_IPOBA STANDARD; PRT; 492 AA.
 AC P07145;
 DT 01-APR-1988 (REL. 07, CREATED)
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE CATALASE (EC 1.11.1.6).
 OS IPOMOEA BATATAS (SWEET POTATO) (BATATE).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC SOLANALES; CONVOLVULACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. KOKEI NO. 14;
 RM 87246622
 RA SAKAJI S., NAKAMURA K., ASAH I T.;
 RL EUR. J. BIOCHEM. 165:437-442(1987).
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN

CC PEROXIDASE.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: HEME GROUP.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
 CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
 DR EMBL; X05549; IBCATR.
 DR PIR; S07124; S07124.
 DR PROSITE; PS00437; CATALASE_1.
 DR PROSITE; PS00438; CATALASE_2.
 KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
 KW PEROXISOME.
 FT ACT_SITE 65 65 BY SIMILARITY.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 492 AA; 56985 MW; 1209173 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                || ||| || |
  CALTVGSRGPILLEDYHLEKIQNFTRRERIPERVVHARGASAKGFFEVDITHLTCAFLRAPGVQTPLIV
  30          40          50          60          70          80          90         100

  RFSTVIHERGSPETIRDPRGFAVKMYTRGGNWDLVGNFP
      110      120      130      140

```

6. US-08-249-182-10 (1-12)

CAT1_LYCES CATALASE ISOZYME 1 (EC 1.11.1.6).

ID CAT1_LYCES STANDARD; PRT; 492 AA.
 AC P30264;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
 DE CATALASE ISOZYME 1 (EC 1.11.1.6).
 GN CAT1.
 OS LYCOPERSICON ESCULENTUM (TOMATO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA DRORY A., WOODSON W.R.;
 RL SUBMITTED (XXX-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
 CC PEROXIDE.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: HEME GROUP.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
 CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
 DR EMBL; M93719; LECAT1A.
 DR PROSITE; PS00437; CATALASE_1.
 DR PROSITE; PS00438; CATALASE_2.
 KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
 KW PEROXISOME; MULTIGENE FAMILY.
 FT ACT_SITE 65 65 BY SIMILARITY.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 492 AA; 56505 MW; 1223422 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10 X
                                     DIEHLTSLDFFR
                                     || ||| || |
SSLTVGPRGPVLLEDYYLIEKLATFDREKIPERVVHARGASAKGFFEVDHSHLTCADFLRAPGAQTPVIC
30      40      50      60      70      80      90      100

RFSTVVHERGSPESIRDIRGFAVKFYTREGNFDLVGNVNP
      110      120      130      140
```

7. US-08-249-182-10 (1-12)

CAT1_GOSHI CATALASE ISOZYME 1 (EC 1.11.1.6).

ID CAT1_GOSHI STANDARD; PRT; 492 AA.
AC P17598;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE CATALASE ISOZYME 1 (EC 1.11.1.6).
GN CAT1.
OS GOSSYPIMUM HIRSUTUM (UPLAND COTTON).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC MALVALES; MALVACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DELTAPINE 62; TISSUE=COTYLEDON;
RM 90304227
RA NI W., TURLEY R.B., TRELEASE R.N.;
RL BIOCHIM. BIOPHYS. ACTA 1049:219-222(1990).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
CC -!- THERE ARE AT LEAST FIVE ISOZYMES OF CATALASE IN COTTON.
CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
DR EMBL; X52135; GHCAT1.
DR PIR; S10770; S10770.
DR PIR; S10395; S10395.
DR PROSITE; PS00437; CATALASE_1.
DR PROSITE; PS00438; CATALASE_2.
KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
KW PEROXISOME; MULTIGENE FAMILY.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 492 AA; 56859 MW; 1209561 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
Residue Identity = 66% Matches = 8 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10 X
                                     DIEHLTSLDFFR
                                     || ||| || |
SSLTVGPRGQVLLEDYHLVEKLANFDRERIPERVVHARGASAKGFFDVTDHSHLTCADFLRAPGVQTPVIV
30      40      50      60      70      80      90      100

RFSTVIHERGSPETLRDPRGFAVKFYTREGNFDLVGNFP
      110      120      130      140
```

8. US-08-249-182-10 (1-12)

CATA_PEA CATALASE (EC 1.11.1.6).

ID CATA_PEA STANDARD; PRT: 494 AA.
 AC P25890;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE CATALASE (EC 1.11.1.6).
 OS PISUM SATIVUM (GARDEN PEA).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
 OC FABACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RM 92032793
 RA ISIN S.H., ALLEN R.D.;
 RL PLANT MOL. BIOL. 17:1263-1265(1991).
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
 CC PEROXIDE.
 CC -!- CATALYTIC ACTIVITY: $2 \text{ H}(2)\text{O}(2) = \text{O}(2) + 2 \text{ H}(2)\text{O}$.
 CC -!- COFACTOR: HEME GROUP.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
 CC -!- SIMILARITY: TO OTHER EUKARYOTIC CATALASES.
 DR EMBL; X60169; PSCATAL.
 DR PIR; S18346; CSPM.
 DR PROSITE; PS00437; CATALASE_1.
 DR PROSITE; PS00438; CATALASE_2.
 KW OXIDOREDUCTASE; PEROXIDASE; IRON; HEME; HYDROGEN PEROXIDE;
 KW PEROXISOME.
 FT ACT_SITE 65 65 BY SIMILARITY.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
 SQ SEQUENCE 494 AA; 57344 MW; 1224857 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.31
 Residue Identity = 66% Matches = 8 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || ||| || |

SSLTVGSRGPILLEDYHLVEKLAQFDRERIPERVVHARGASAKGFFEVTHDISHLTCADFLRAPGVQTPVIV
 30 40 50 60 70 80 90 100

RFSTVIHERGSPETLRDPRGFAVKFYTREGRNYDLVGNNFP
 110 120 130 140

9. US-08-249-182-10 (1-12)

VG74_BPML5 GENE 74 PROTEIN (GP74).

ID VG74_BPML5 STANDARD; PRT: 72 AA.
 AC 005289;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE GENE 74 PROTEIN (GP74).
 GN 74.
 OS MYCOBACTERIOPHAGE L5.
 OC VIRIDAE; NOT YET CLASSIFIED.
 RN [1]

SEQUENCE FROM N.A.
 RM 93211282
 RA HATFULL G.F., SARKIS G.J.;
 RL MOL. MICROBIOL. 7:395-405(1993).
 DR EMBL; Z18946; MLCGA.
 SQ SEQUENCE 72 AA; 8469 MW; 24914 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.25
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

      X      10 X
      DIEHLTSLDFFR
      |||||
MEFPFGGTMFQCLHLTSLDKVQLWFRWRSRDWQFGHAKPIPEGMDAASVANLMRQFIVGEISHDEYNRLLPN
      10      20 X      30      40      50      60      70

```

10. US-08-249-182-10 (1-12)
 YSCA_BACSU HYPOTHETICAL 13.6 KD PROTEIN IN SECA 5'REGION (ORF)

ID YSCA_BACSU STANDARD; PRT; 117 AA.
 AC P28368;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 13.6 KD PROTEIN IN SECA 5'REGION (ORF1).
 OS BACILLUS SUBTILIS.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91192600
 RA SADAIE Y., TAKAMATSU H., YAMANE K.;
 RL GENE 98:101-105(1991).
 DR EMBL; D90218; BSSECA.
 DR EMBL; D10279; BSSECA1.
 DR PIR; JN0145; JN0145.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 117 AA; 13593 MW; 67264 CN;

Initial Score = 6 Optimized Score = 7 Significance = 4.25
 Residue Identity = 58% Matches = 7 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||| ||| |
YNAIDLATNKLKLRQIRKHKTKVNRKFREQGSPKYLLANGLGSDTDIAVQDDIEEEESLDIVRQKRFLNKPMD
      10      20      30      40      50 X      60 X      70

SEEAILQMNMLGHNFFVFTNAETNLTNVVYRRNDGKYGLI
      80      90      100      110

```

11. US-08-249-182-10 (1-12)
 VG04_VARV PROTEIN G4.

ID VG04_VARV STANDARD; PRT; 124 AA.
 AC P32994;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE PROTEIN G4.
 GN G4L OR H4L.
 OS VARIOLA VIRUS.
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;

CC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=INDIA-1967 / ISOLATE IND3;
 RA SHCHELKUNOV S.N., BLINOV V.M., RESENCHUK S.M., TOTMENIN A.V.,
 RA SANDAKHCHIEV L.S.;
 RL VIRUS RES. 30:239-258(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=INDIA-1967 / ISOLATE IND3;
 RM 93190624
 RA SHCHELKUNOV S.N., BLINOV V.M., TOTMENIN A.V., MARENNIKOVA S.S.,
 RA KOLYKHALOV A.A., FROLOV I.V., CHIZHIKOV V.E., GYTOROV V.V.,
 RA GASHIKOV P.V., BELANOV E.F., BELAVIN P.A., RESENCHUK S.M.,
 RA ANDZHAPARIDZE D.G., SANDAKHCHIEV L.S.;
 RL VIRUS RES. 27:25-35(1993).
 RN [3]
 RP COMPLETE GENOME.
 RC STRAIN=INDIA-1967 / ISOLATE IND3;
 RA BLINOV V.M.;
 RL SUBMITTED (NOV-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; X67119; VVHINDQKH.
 DR EMBL; X69198; VVCGAA.
 DR PIR; A36844; A36844.
 DR PIR; S33081; S33081.
 SQ SEQUENCE 124 AA; 14014 MW; 86978 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.25
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                X      10 X
                DIEHLTSLDFFR
                || |  | ||
MKNVLIIFGKPYCSICENVSEAVEELKSEYDILHVDILSFFLKDGSSMLGDVKRGTLIGNFAAHL SNYIVS
      10      20      30      40 X      50      60      70

IFKYNPQTKQMAFVDINKSL
      80      90
  
```

12. US-08-249-182-10 (1-12)
 VG04_VACCC PROTEIN G4.

ID VG04_VACCC STANDARD; PRT; 124 AA.
 AC P21025;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE PROTEIN G4.
 GN G4L.
 OS VACCINIA VIRUS (STRAIN COPENHAGEN).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91021027
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
 RA PADLETTI E.;
 RL VIROLOGY 179:247-266(1990).
 RN [2]
 RP COMPLETE GENOME.
 RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
 RA PADLETTI E.;
 RL VIROLOGY 179:517-563(1990).
 DR EMBL; M35027; PXVACCG.

DR PIR; 142511; 142511;
SQ SEQUENCE 124 AA; 13987 MW; 88525 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.25
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```

              X      10 X
              DIEHLTSLDFFR
              || | | ||
MKNVLIIFGKPYCSICENVSDAVEELKSEYDILHVDILSFFLKDGDSMLGDVKRGTIGNFAAHLSNYIVS
      10      20      30      40 X      50      60      70

IFKYNPQTKQMAFVDINKSL
      80      90
```

13. US-08-249-182-10 (1-12)

YME1_BACSU HYPOTHETICAL 17.2 KD PROTEIN IN MEND 5'REGION (ORF)

ID YME1_BACSU STANDARD; PRT; 145 AA.
AC P23972;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 17.2 KD PROTEIN IN MEND 5'REGION (ORF1).
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA ROWLAND B., HILL K., MUELLER J., DRISCOLL J., TABER H.;
RL SUBMITTED (OCT-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; M74538; BSMENAQOP.
DR PIR; S27507; S27507.
KW HYPOTHETICAL PROTEIN; MENAQUINONE BIOSYNTHESIS.
SQ SEQUENCE 145 AA; 17159 MW; 108895 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.25
Residue Identity = 50% Matches = 6 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

```

              X      10 X
              DIEHLTSLDFFR
              || | | ||
MVTTVQRTFRKKVLHALHKAKEVNHAVLISYSRQIESLDPLSFFNYGAKKYTGNRFFWSDPESELTIVGLGK
      10      20      30 X      40 X      50      60      70

EAVFQTNQKNSEYREVFEQWER
      80      90
```

14. US-08-249-182-10 (1-12)

YKAB_CAEEL HYPOTHETICAL 27.3 KD PROTEIN B0303.8 IN CHROMOSOME

ID YKAB_CAEEL STANDARD; PRT; 239 AA.
AC P34259;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.3 KD PROTEIN B0303.8 IN CHROMOSOME III.
GN B0303.8.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELEMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;

RA SULSTON J., DU Z., THOMAS K., WILSON R., HILLIER L., STADEN R.,
 RA HALLORAN N., GREEN P., THIERRY-MIEG J., QIU L., DEAR S., COULSON A.,
 RA CRAXTON M., DURBIN R.K., BERKS M., METZSTEIN M., HAWKINS T.,
 RA AINSCOUGH R., WATERSTON R.;
 RL NATURE 356:37-41(1992).
 DR EMBL; M77697; CEB0303.
 DR PIR; S27789; S27789.
 DR WORMPEP; B0303.8; CE00562.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 239 AA; 27341 MW; 323313 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.25
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 DIEHLTSLDFFR
 || || ||
 SVRSTFYCENGYFGPICDRRSRTFAPKSDIQSTSTPGYQTQVLKFDKISDDIIYSSLAFFVLLLIIFNCIL
 120 130 140 150 160 X 170 X 180
 CCYRPKKSQKYL DVSLGSPKVF SICGYSADKSGNTTEYLD
 190 200 210 220

15. US-08-249-182-10 (1-12)

APPY_ECOLI APPY PROTEIN (M5 POLYPEPTIDE).

ID APPY_ECOLI STANDARD; PRT; 243 AA.
 AC P05052;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)
 DE APPY PROTEIN (M5 POLYPEPTIDE).
 GN APPY.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 89155479
 RA ATLUNG T., NIELSEN A., HANSEN F.G.;
 RL J. BACTERIOL. 171:1683-1691(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JM 83;
 RM 87231084
 RA KEMP E.H., MINTON N.P., MANN N.H.;
 RL NUCLEIC ACIDS RES. 15:3924-3924(1987).
 CC -!- FUNCTION: INDUCES THE SYNTHESIS OF ACID PHOSPHATASE (APPA) AND
 CC SEVERAL OTHER POLYPEPTIDES DURING THE DECELERATION PHASE OF
 CC GROWTH. IT ALSO ACTS AS A TRANSCRIPTIONAL REPRESSOR FOR ONE GROUP
 CC OF PROTEINS THAT ARE SYNTHESIZED PREFERENTIALLY IN EXPONENTIAL
 CC GROWTH AND FOR ONE GROUP SYNTHESIZED ONLY IN THE STATIONARY PHASE.
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; M24530; ECAPPYAA.
 DR EMBL; Y00138; ECM5.
 DR PIR; A29260; BVECM5.
 DR PIR; JS0110; BVECAV.
 DR ECGENE; EG10050; APPY.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY.
 KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
 FT DNA_BIND 149 168 H-T-H MOTIF (BY SIMILARITY).
 FT CONFLICT 25 25 F -> FKKNSLF (IN REF. 2).

Initial Score = 6 Optimized Score = 6 Significance = 4.25
 Residue Identity = 50% Matches = 6 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10 X
                                DIEHLTSLDFFR
                                ||  || ||
MDYVCSVVFICQSFDLIINRRVISFIVSDKIRRELPVCPSKLRIVDIDKKTCLSFFIDVNNELPGKFTLDKN
      10      20      30      40      X 50      X 60      70

```

```

GYIAEEEEPPLSLVFSLFEGIKIADSHSLWLERLC
      80      90     100

```

> 0 <
 0| 0 IntelliGenetics
 > 0 <

seq. 11

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_11a.res made by on Thu 22 Sep 94 10:36:47-PDT.

Query sequence being compared: US-08-249-182-11 (1-23)
 Number of sequences searched: 42145
 Number of scores above cutoff: 4803

Results of the initial comparison of US-08-249-182-11 (1-23) with:
 Data bank : A-GeneSeq 15, all entries

100000-

N -

U50000-

M -

B -

E -

R - *

O *

F10000-

S -

E 5000- * *

Q -

U -

E -

N - *

C -

E -

S 1000-

500- *

100-

50-

10-
-
-
5-
-
-
-
-
-
0

SCORE 0 | 2 | 3 | 5 | 7 | 10 | 12 | 15 | 17 | 20 | 22
STDEV 0 2 3 4 5 6 7 8 9

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.52

```
Times:                CPU                Total Elapsed
                00:00:29.95                00:00:39.00
```

```
Number of residues:      5287517
Number of sequences searched: 42145
Number of scores above cutoff: 4803
```

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.
Cut-off raised to 5.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37453	**** 13 standard deviations above mean **** Autotaxin peptide ATX 103.	23	22	22	13.81	0
2. P82098	**** 5 standard deviations above mean **** Protein 2 encoded by anylolylt	33	9	9	5.26	0
	**** 4 standard deviations above mean ****					

3. R23482	glyn.	477	8	8	4.60	0
4. R06262	Calf acetylcholine receptor (520	8	8	4.60	0
5. R06261	Mouse acetylcholine receptor	520	8	8	4.60	0
6. R33646	Rabbit pre-pro serum albumin.	608	8	8	4.60	0
7. R11255	Murine IL-4 receptor.	810	8	8	4.60	0
8. R04574	Derived amino acid sequence o	810	8	9	4.60	0
9. R11993	Glutamate receptor 5-1.	920	8	8	4.60	0
10. R32882	Cardiac adenylyl cyclase type	1184	8	8	4.60	0
11. R40227	ACVS.	3639	8	8	4.60	0
12. R13896	ACV synthetase.	3712	8	8	4.60	0
**** 3 standard deviations above mean ****						
13. P90430	Cyclophilin.	42	7	8	3.95	0
14. R32203	Apple fruit PPD pSR7.	68	7	7	3.95	0
15. R25119	Non-A, Non-B Hepatitis Virus	72	7	8	3.95	0
16. P40026	Fusion protein of insulin-lik	89	7	7	3.95	0
17. R12603	SIB 121 intestinal mucin.	95	7	7	3.95	0
18. P60578	Human prepro-somatostatin-C.	119	7	7	3.95	0
19. P70277	Sequence of pre-pro-insulin-1	195	7	7	3.95	0
20. R12521	B cell differentiation factor	212	7	9	3.95	0
21. P90363	Recombinant Group C Eimeria t	226	7	8	3.95	0
22. R05026	Beta subunit of rat high affi	243	7	7	3.95	0
23. R14770	Beta subunit of high affinity	246	7	7	3.95	0
24. R22997	Yeast proteasome YC7-alpha su	252	7	7	3.95	0
25. P71672	Human serine protease.	262	7	7	3.95	0
26. R06639	Orotidine-5'-phosphate decarb	267	7	7	3.95	0
27. R06640	Orotidine-5'-phosphate decarb	268	7	7	3.95	0
28. R13467	Cc protein.	276	7	8	3.95	0
29. R03339	VP1 sequence for HRV serotype	287	7	9	3.95	0
30. R03340	VP1 sequence for HRV serotype	291	7	8	3.95	0
31. P60326	Interleukin-1 gene product.	295	7	7	3.95	0
32. R37346	PEP PM.	313	7	7	3.95	0
33. P93143	Sequence encoded by ORF2 of p	314	7	7	3.95	0
34. R33279	43 kD endoflagellum sheath pr	320	7	9	3.95	0
35. R25116	Non-A, Non-B Hepatitis Virus	336	7	9	3.95	0
36. R27163	CD2 binding LFA-3-Ig fusion p	347	7	7	3.95	0
37. R42424	Rat gustducin alpha subunit.	354	7	7	3.95	0
38. R25698	Murine adrenergic beta-3 rece	388	7	7	3.95	0
39. R07130	H20B receptor.	392	7	7	3.95	0
40. R32501	Beta-adrenergic receptor.	400	7	7	3.95	0

1. US-08-249-182-11 (1-23)

R37453 Autotaxin peptide ATX 103.

ID R37453 standard; peptide; 23 AA.
AC R37453;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 103.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 103. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body

CC Fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.

SQ Sequence 23 AA;
SQ 0 A; 1 R; 2 N; 2 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 2 G; 0 H;
SQ 1 I; 4 L; 0 K; 0 M; 1 F; 1 P; 1 S; 4 T; 0 W; 1 Y; 2 V;

Initial Score = 22 Optimized Score = 22 Significance = 13.81
Residue Identity = 95% Matches = 22 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

X      10      20 X
TEFLSNYLTNVDDITLVPETLGR
|||||
TEFLSNYLTNVDDITLVPETLGR
X      10      20 X
```

2. US-08-249-182-11 (1-23)

P82098 Protein 2 encoded by amylolytic enzyme expression/

ID P82098 standard; protein; 33 AA.
AC P82098;
DT 29-OCT-1990 (first entry)
DE Protein 2 encoded by amylolytic enzyme expression/secretion sequence
KW amylolytic enzyme; protein secretion; ss.
OS synthetic.
PN J63219381-A.
PD 13-SEP-1988.
PF 09-MAR-1987; 052072.
PR 09-MAR-1987; JP-052072.
PA (OJIP) OJI Paper KK; (OJIC-) OJI Corn Starch KK.
DR WPI; 88-297738/42.
PT DNA contg regions relating to expression and secretion of protein -
PT comprise promoter region relating to amylolytic enzyme gene,
PT ribosome region, region relating to secretion of protein expression
PT product
PS Claim 6; page 510; 14pp; Japanese.
CC DNA includes a promoter region for expression of the amylolytic
CC enzyme of an alkali Bacillus, the ribosome binding site and the
CC region responsible for secretion of the enzyme.
CC See also N82034, P82097 and N82033.
SQ Sequence 33 AA;
SQ 2 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 2 G; 0 H;
SQ 3 I; 6 L; 3 K; 2 M; 3 F; 1 P; 2 S; 3 T; 0 W; 0 Y; 3 V;

Initial Score = 9 Optimized Score = 9 Significance = 5.26
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

```

X      10      20 X
TEFLSNYLTNVDDITLVPETLGR
||| | || | ||
MKMRTGKKGFLSILLAFLLVITSIPFTLV DVEA
10      20      30
```

3. US-08-249-182-11 (1-23)

R25462 glgA.

ID R25462 standard; Protein; 477 AA.
AC R25462;
DT 15-JAN-1993 (first entry)
DE glgA.
KW Glycogen synthase gene; glgA; polymerase chain reaction; k12 618;

KW plastid, specific gravity, free sugar content.
 OS Escherichia coli.
 PN W09211382-A.
 PD 09-JUL-1992.
 PF 19-DEC-1991; U09654.
 PR 21-DEC-1990; US-632383.
 PR 16-JUL-1991; US-731226.
 PR 24-JUL-1991; US-735065.
 PA (CALJ) CALGENE INC.
 PI Shewmaker CK, Stalker DM;
 DR WPI; 92-250101/30.
 DR N-PSDB; 025978.
 PT Glycogen biosynthesis enzyme-encoding genes - for prodn. of
 PT transgenic plants having modified starch content, e.g. decreased
 PT amylase
 PS Disclosure; Fig 2; 65pp; English.
 CC The sequence given is encoded by the E. coli glycogen synthase gene
 CC (glgA) which was generated by polymerase chain reaction from E. coli
 CC strain k12 618. The sequence encoding this polypeptide was used in
 CC the construction of an oligonucleotide which also contained an
 CC endogenous plant sequence. This construct is useful in studying and
 CC manipulating the starch biosynthetic pathway. This enzyme can be
 CC targeted to a plastid where starch synthesis occurs. Plants or plnt
 CC part which synthesise and store starch may be obtained which have
 CC increased or decreased starch content and modified starch related
 CC properties, such as specific gravity, free sugar content and/or
 CC novel and useful starches, eg. potato starch with decrease amylose and
 CC modified amylopectin.
 SQ Sequence 477 AA;
 SQ 51 A; 30 R; 11 N; 29 D; 0 B; 3 C; 21 Q; 20 E; 0 Z; 45 G; 16 H;
 SQ 17 I; 60 L; 12 K; 10 M; 23 F; 23 P; 25 S; 18 T; 9 W; 20 Y; 34 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

					X	10	20
					TEFLSNYL	TNVD	ITLVPETLG
GLEFNGQISFLKAGLYYADHITAVSPTYAREITEPQFAYGMEGLLQQRHREGRLSGVLNGVDEKIWSPETDL							
190	200	210	220	230	240	250	
X							
R							
LLASRYTRDTLEDKAENKRQLQIAMGLKVDDKVPLFAVVSRLTSQKGLDLV							
260	270	280	290	300	310		

4. US-08-249-182-11 (1-23)
 R06262 Calf acetylcholine receptor (AChR) delta-subunit.

ID R06262 standard; protein; 520 AA.
 AC R06262;
 DT 07-DEC-1990 (first entry)
 DE Calf acetylcholine receptor (AChR) delta-subunit.
 KW Nicotinic acetyl choline receptor; AchR; TE671; insecticides;
 KW Muscle relaxants; anthelmintics;
 OS Bos Taurus.
 PN CA2003459-A.
 PD 23-MAY-1990.
 PF 21-NOV-1989; 003459.
 PR 23-NOV-1988; US-275422.
 PA (SALK) SALK INST FOR BIOL STUD.
 PI Lindstrom JM, Schoepfer RD;
 DR WPI; 90-231525/31.

Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYL TNVDDITLVPETLG
                                ||  ||  ||  ||
  LTLGLLAALVVCALPGSWGLNEEQRLIQHLFNEKGYDKDLRPVARKEDKVDVALSLTSLNLISLKEVEETLT
    10      20      30      40      50      X 60      70
  
```

X
R

```

  TNVWIDHAWVDSRLQWDANDFGNITYLALPPDMVWLPEIVLENNNDGSFQI
    80      90      100      110      120
  
```

6. US-08-249-182-11 (1-23)

R33646 Rabbit pre-pro serum albumin.

ID R33646 standard; Protein; 608 AA.
 AC R33646;
 DT 09-JUL-1993 (first entry)
 DE Rabbit pre-pro serum albumin.
 KW RSA; large scale prepn.; culture medium; medicine carrier.
 OS *Oryctolagus cuniculus*.
 PN J05038287-A.
 PD 19-FEB-1993.
 PF 10-JUL-1991; 194984.
 PR 10-JUL-1991; JP-194984.
 PA (TOYJ) TOSOH CORP.
 DR WPI; 93-096421/12.
 DR N-PSDB; 038280.
 PT Rabbit serum albumin derived from pre pro-serum albumin gene -
 PT comprises pre pro-albumin genes expressed by defined
 PT poly:peptide and DNA sequences
 PS Disclosure; Fig 2; 19pp; Japanese.
 CC The sequence is that of the rabbit pre-pro serum albumin which
 CC may be recombinantly produced on a large scale in e.g. *Pichia pastoris*.
 CC The rabbit serum albumin (RSA) may be used for the prepn. of a
 CC culture medium of animal cells and as a carrier for medicines.
 SQ Sequence 608 AA;
 SQ 56 A; 25 R; 12 N; 43 D; 0 B; 35 C; 14 Q; 56 E; 0 Z; 20 G; 23 H;
 SQ 17 I; 65 L; 59 K; 2 M; 28 F; 29 P; 29 S; 28 T; 2 W; 25 Y; 40 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYL TNVDDITLVPETLG
                                ||  ||  ||  ||
  EDYLSVVLNRLCVLHEKTPVSEKVTKCCSESLVDRRPFCSALGPDETVVPKEFNAETFTFHADICTLPETER
    480      490      500      510      520      X 530      540
  
```

X
R

```

  KIKKQ TALVELVKHKPHATNDQLKTVVGFE TALLDKCCSAEDKEACFAVEG
  X 550      560      570      580      590
  
```

7. US-08-249-182-11 (1-23)

R11255 Murine IL-4 receptor.

ID R11255 standard; Protein; 810 AA.
 AC R11255;

PT the effects of agents which affect acetylcholine receptors in
 PT skeletal muscles.
 PS Disclosure; p; English.
 CC Receptors may be used in assay for materials which modify them.
 CC They may be produced in substantial, pure quantities for use in
 CC experimentation, development of insecticides without effect on
 CC hMNARs and treatment of parasitic infections. MABs raised to the
 CC peptides may be useful in detection of the structure of MNARs.
 CC 24 unidentified residues are due to the poor quality of the
 CC sequence reproduction.
 SQ Sequence 520 AA;
 SQ 29 A; 24 R; 27 N; 25 D; 0 B; 9 C; 17 Q; 33 E; 0 Z; 22 G; 9 H;
 SQ 34 I; 61 L; 23 K; 9 M; 27 F; 33 P; 39 S; 22 T; 15 W; 15 Y; 47 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTNVDDITLVPETLG
 || || | |||
 LTLGLLAALVVCALPGSWG LNEEQRLIQHLFNEKGYDKDLRPVARKEDKVDVALSLTSLNLISLKEVEETLT
 10 20 30 40 50 X 60 70

 X
 R

 TNVWIDHAWVDSRLQWDANDFGNITVLALPPDMVWLPEIVLENNNDGSFQI
 80 90 100 110 120

5. US-08-249-182-11 (1-23)

R06261 Mouse acetylcholine receptor (AChR) delta-subunit.

ID R06261 standard; protein; 520 AA.
 AC R06261;
 DT 07-DEC-1990 (first entry)
 DE Mouse acetylcholine receptor (AChR) delta-subunit.
 KW Nicotinic acetyl choline receptor; AChR; TE671; insecticides;
 KW Muscle relaxants; anthelmintics;
 OS Mus sp.
 PN CA2003459-A.
 PD 23-MAY-1990.
 PF 21-NOV-1989; 003459.
 PR 23-NOV-1988; US-275422.
 PA (SALK) SALK INST FOR BIOL STUD.
 PI Lindstrom JM, Schoepfer RD;
 DR WPI; 90-231525/31.
 PT Human muscle nicotinic acetylcholine receptor - used to assay
 PT the effects of agents which affect acetylcholine receptors in
 PT skeletal muscles.
 PS Disclosure; p; English.
 CC Receptors may be used in assay for materials which modify them.
 CC They may be produced in substantial, pure quantities for use in
 CC experimentation, development of insecticides without effect on
 CC hMNARs and treatment of parasitic infections. MABs raised to the
 CC peptides may be useful in detection of the structure of MNARs.
 CC Unidentified residues are due to the poor quality of the
 CC sequence reproduction.
 SQ Sequence 520 AA;
 SQ 28 A; 25 R; 28 N; 25 D; 0 B; 9 C; 17 Q; 33 E; 0 Z; 23 G; 9 H;
 SQ 32 I; 61 L; 22 K; 12 M; 27 F; 33 P; 39 S; 24 T; 10 W; 16 Y; 46 V;
 SQ 1 Others;

Initial Score = 8 Optimized Score = 8 Significance = 4.60

DT 30-MAY-1991 (first entry)
 DE Murine IL-4 receptor.
 KW Interleukin-4; soluble; IgE.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Region 1..25
 FT /label= signal sequence
 FT Domain 26..233
 FT /label= extracellular domain
 FT Domain 234..257
 FT /label= transmembrane region
 FT Domain 258..810
 FT /label= cytoplasmic domain
 PN EP-419091-A.
 PD 27-MAR-1991.
 PF 05-SEP-1990; 309700.
 PR 07-SEP-1989; US-404179.
 PR 20-MAR-1990; US-496449.
 PA (SCHE) SCHERING CORP.
 PI Galizzi JP, Harada N, Miyajima A;
 DR WPI; 91-088700/13.
 PT Nucleic acid encoding mammalian interleukin-4 receptor - used as
 PT antagonists of interleukin-4 in treating conditions associated
 PT with excess IgE prodn. including allergic conditions.
 PS Disclosure; Fig 5; 18pp; English.
 CC The sequence was deduced from cDNA clone 19 from a cDNA
 CC library prepd. from MC/9 cells.
 CC See also R11254 and Q11055.
 SQ Sequence 810 AA;
 SQ 50 A; 23 R; 29 N; 34 D; 0 B; 31 C; 39 Q; 53 E; 0 Z; 63 G; 18 H;
 SQ 29 I; 81 L; 30 K; 15 M; 23 F; 82 P; 90 S; 38 T; 17 W; 18 Y; 47 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTVDDITLVPETLG
 || || || ||
 KAFSSLLSSNGIRGDTAAAGTDDGHGGYKPFQNPVPNQSPSSVPLFTFGLDSELSPLNSDPPKSPPECLG
 610 620 630 640 650 X 660 670

X
 R
 LELGLKGGDWVKAPPPADQVPKPFQDGLGFGIVYSSLTCHLCGHLKQHHSQ
 X 680 690 700 710 720

8. US-08-249-182-11 (1-23)

R04574 Derived amino acid sequence of coding region of nu

ID R04574 standard; protein; 810 AA.
 AC R04574;
 DT 17-SEP-1990 (first entry)
 DE Derived amino acid sequence of coding region of murine IL-4 receptor
 KW mammalian interleukin-4 receptor; cytokine; antibody production;
 OS synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 209..232
 FT /label= putative transmembrane region
 PN EP-367566-A.
 PD 09-MAY-1990.
 PF 31-OCT-1989; 311244.
 PR 23-JUN-1989; US-370924.
 PA (IMMU-) Immunex Corp.

PT COSMAN D; PARK L; HOSLEY B; BECKMANN P; HARTH CO; IBERDA R;
 DR WPI; 90-141470/19.
 DR N-PSDB; 004305.
 PT Recombinant mammalian interleukin-4 receptor used in diagnosis,
 PT assays and therapy and for prodn. of antibodies for diagnosis,therapy
 PT and for prodn. of antibodies
 PS Disclosure; p; English.
 CC The interleukin-4 receptor can be used to regulate immune responses or to
 CC treat IgE-induced hypersensitivity.
 CC See also 004307.
 SQ Sequence 810 AA;
 SQ 49 A; 23 R; 28 N; 34 D; 0 B; 32 C; 39 Q; 52 E; 0 Z; 63 G; 18 H;
 SQ 29 I; 80 L; 30 K; 15 M; 25 F; 83 P; 89 S; 39 T; 16 W; 18 Y; 48 V;

Initial Score = 8 Optimized Score = 9 Significance = 4.60
 Residue Identity = 43% Matches = 10 Mismatches = 12
 Gaps = 1 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTNVDDITLVPETLG
 || || | | || ||
 AFSSLLSSNAIRGDTAAAGTDDGHHGGYKPFQNPVPNQSPSSVPLFTFGLDTE-LSPSPNSDPPKSPPECLG
 610 620 630 640 650 X 660 670

X
 R

LELGLKGGDWVKAPPPADEVPKPFQDDLGFGIVYSSLTCHLCGHLKQHHSSQ
 X 680 690 700 710 720

9. US-08-249-182-11 (1-23)

R11993 Glutamate receptor 5-1.

ID R11993 standard; Protein; 920 AA.
 AC R11993;
 DT 31-JUL-1991 (first entry)
 DE Glutamate receptor 5-1.
 KW Glutamate receptor 5-1; probe; ligand; drug screening.
 OS Rattus rattus.
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT /label= sig_peptide
 FT Protein 31..920
 FT /label= mat_protein
 FT Region 402..416
 FT /label= insertion
 PN WD9106648-A.
 PD 16-MAY-1991.
 PF 25-OCT-1990; U06153.
 PR 27-OCT-1989; US-428116.
 PA (SALK) SALK INST FOR BIOL STUD.
 PI Heinemann SF, Boulter JR, Hollmann M, Bettler B, Jensen JE;
 DR WPI; 91-164197/22.
 DR N-PSDB; 011853.
 PT Glutamate receptors - used to screen for functional ligands and
 PT identify and isolate further receptors
 PS Disclosure; Fig 10; 109pp; English.
 CC GluR5-1 has a 15 amino acid insert (see features) compared to the
 CC shorter variant GluR5-2 and is unique among the receptors GluR1-7.
 CC It has a Mr of 100,000. The signal sequence cleavage site is after
 CC a Pro, which is atypical.
 CC The gene and protein can be used in drug screening, to
 CC determine whether a substance is a functional ligand for the
 CC receptor by monitoring ion channel activity.
 CC See also 011849-855.

Sequence 920 AA;
SQ 53 A; 47 R; 45 N; 46 D; 0 B; 13 C; 28 Q; 54 E; 0 Z; 55 G; 14 H;
SQ 63 I; 99 L; 59 K; 28 M; 40 F; 37 P; 70 S; 59 T; 18 W; 38 Y; 54 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10      20
                                     TEFLSNYLTNVDDITLVPETLG
                                     ||| | | |||
NRDRSNNITDSLANTRLIVTTILEEPYVMYRKSDKPLYGNDRFEGYCLDLLKELSNILGFLYDVKLVPDGKY
      440      450      460      470      480 X      490      500
```

X
R

```
GAQNDKGEWNGMVKELIDHRADLAVAPLTITYVREKVIDFSKPFMTLGISI
X 510      520      530      540      550
```

10. US-08-249-182-11 (1-23)

R32882 Cardiac adenylyl cyclase type V.

ID R32882 standard; Protein; 1184 AA.
AC R32882;
DT 17-JUN-1993 (first entry)
DE Cardiac adenylyl cyclase type V.
KW CACV; therapy; diagnostic; cardiac function; cyclic AMP; cAMP; heart
KW failure.
OS Canis familiaris.
PN EP-529622-A.
PD 03-MAR-1993.
PF 27-AUG-1992; 114637.
PR 29-AUG-1991; US-751460.
PA (AMCY) AMERICAN CYANAMID CO.
PI Ishikawa Y, Konski AF;
DR WPI; 93-068688/09.
DR N-PSDB; 037543.
PT Isolated nucleic acid mol. encoding Cardiac adenylyl cyclase type
PT V - useful for determining and modifying cardiac function
PS Claim 4; Page 15-27; 38pp; English.
CC Left ventricular tissue of canine heart was used as a source of mRNA.
CC A cDNA library was prepd. in lambda gt10 phage. A 970 bp Aat-HincII
CC fragment from type I adenylyl cyclase cDNA was used as probe. The
CC clones isolated were used to obtain cDNA encoding CACV. This probe
CC may also be used to screen a human cardiac cDNA library to obtain
CC the cDNA encoding human CACV. CACV, its analogues and antibodies
CC are useful in therapy or diagnostic assays, e.g. in modifying and
CC determining cardiac function. A decrease in CACV content of the
CC heart contributes to impaired cAMP prodn. and in heart failure. The
CC CACV can also be used to screen for cpds. which stimulate or inhibit
CC the activity of the cyclase.
SQ Sequence 1184 AA;
SQ 123A; 83 R; 52 N; 43 D; 0 B; 35 C; 43 Q; 64 E; 0 Z; 84 G; 30 H;
SQ 58 I; 122L; 47 K; 36 M; 54 F; 51 P; 71 S; 52 T; 11 W; 32 Y; 93 V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10      20
                                     TEFLSNYLTNVDDITLVPETLG
                                     || ||| |||
VRKFLTLTFREPDLKKYSKQVDDRFQAYVACASLVFLFCFVQITIVPHSVFMLSFYLTCLLLTLVVFVSV
      660      670      680      690      700      710      720      730
```

X
R

IYSCVKLFPGLQLSRKIVRSKTNSTLVGVFTITLVFLSAFVNMFCNSE
X 740 750 760 770 780

11. US-08-249-182-11 (1-23)
R40227 ACVS.

ID R40227 standard; Protein; 3639 AA.
AC R40227;
DT 21-FEB-1994 (first entry)
DE ACVS.
KW Delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase; ACVS;
KW beta-lactam; antibiotic; transformed; cephalosporin; vector.
OS Acremonium chrysogenum.
PN J05192162-A.
PD 03-AUG-1993.
PF 25-JUL-1991; 186222.
PR 31-JUL-1990; JP-205677.
PA (TAKE) TAKEDA CHEM IND LTD.
DR WPI; 93-277475/35.
DR N-PSDB; 048231.
PT DNA coding delta-(L-alpha-amino-adipyl) L-cystinyl D-valine
PT synthase - for improved productivity of cephalosporin antibiotics
PS Claim 1; Page 14-27; 69pp; Japanese.
CC The sequence (048231) is of a vector which includes the
CC delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase gene.
CC This sequence was transformed into a host cell to express the ACVS
CC product. The protein produced (R40227) was then used to manufacture
CC a beta-lactam antibiotic.
SQ Sequence 3639 AA;
SQ 258A; 209R; 143N; 213D; 0 B; 39 C; 175Q; 232E; 0 Z; 232G; 126H;
SQ 188I; 405L; 135K; 58 M; 130F; 181P; 271S; 212T; 38 W; 127Y; 267V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTNVDDITLVPETLG
||||| | ||
TQKPSDLAYVIFTSGTTGKPKGVLEHQSIVVQLRNSLIERYFGETNGSHAVLFLSNVVFDFSLEQLCLSVLG
2510 2520 2530 2540 2550 2560 2570

X
R

GNKLIIPPEEGLTHEAFYDIGRREKLSYLSGTPSVLQIELSRPLHLMVT
2580 2590 2600 2610 2620

12. US-08-249-182-11 (1-23)
R13896 ACV synthetase.

ID R13896 standard; Protein; 3712 AA.
AC R13896;
DT 22-NOV-1991 (first entry)
DE ACV synthetase.
KW Beta lactam antibiotics; penicillin.
OS Acremonium chrysogenum.
FH Key Location/Qualifiers
FT Domain 301..1068
FT /label= 1

FT /function= activation of amino acid substrate
 FT Region 374..423
 FT /label= subdomain
 FT Region 474..501
 FT /label= subdomain
 FT Region 655..699
 FT /label= subdomain
 FT Region 725..754
 FT /label= subdomain
 FT Domain 1392..2154
 FT /label= II
 FT /function= activation of amino acid substrate
 FT Region 1470..1518
 FT /label= subdomain
 FT Region 1564..1590
 FT /label= subdomain
 FT Region 1745..1789
 FT /label= subdomain
 FT Region 1817..1846
 FT /label= subdomain
 FT Domain 2474..3295
 FT /label= III
 FT /function= activation of amino acid substrate
 FT Region 2554..2603
 FT /label= subdomain
 FT Region 2647..2673
 FT /label= subdomain
 FT Region 2827..2871
 FT /label= subdomain
 FT Region 2899..2928
 FT /label= subdomain
 FT Domain 3560..3647
 FT /label= IV
 FT /function= thioesterase
 PN EP-445868-A.
 PD 11-SEP-1991.
 PF 27-FEB-1991; 200423.
 PR 28-FEB-1990; EP-200475.
 PR 28-FEB-1990; EP-200488.
 PR 02-JUL-1990; EP-201768.
 PR 03-OCT-1990; EP-202628.
 PR 27-FEB-1991; EP-200423.
 PA (KONN) GIST-BROCADES NV.
 PI Veenstra AE, Martin JF, Garcia BD, Gutierrez S, Barredo JL;
 PI Montenegro PE, Von Doehren H, Palissa H, Van Liempt H;
 DR WPI; 91-268735/37.
 DR N-PSDB; Q13608.
 PT DNA encoding amino:adipyl-cysteinyl-valine synthetase - used for
 PT prodn. of the enzyme or enhanced prodn. of new or known
 PT beta-lactam antibiotic cpds.
 PS Claim 1; Page 20; 54pp; English.
 CC The DNA sequence was obt'd. from five subclones isolated from a
 CC gene library of *A. chrysogenum* C10 (ATCC 48). The protein
 CC sequence was deduced from the DNA. Three distinct regions of
 CC homology have been identified, domains I, II and III. Within
 CC these domains several even more conserved elements can be
 CC distinguished. Since the enzyme synthesises a tripeptide, which
 CC most probably requires the activation of three amino acids, a
 CC role for these domains in the activation reactions seems likely.
 CC A fourth domain is thought to act as a thioesterase.
 CC The gene can be used to express the synthetase enzyme which can
 CC be used for the prodn. of new beta-lactam antibiotics.
 CC See also R13896.
 SQ Sequence 3712 AA;
 SQ 262A; 214R; 144N; 219D; 0 B; 42 C; 180Q; 237E; 0 Z; 237G; 127H;
 SQ 190I; 413L; 135K; 58 M; 131F; 184P; 282S; 217T; 39 W; 128Y; 273V;

Initial Score = 8 Optimized Score = 8 Significance = 4.60
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYLTNVDDITLVPETLG
                                |||||      |      ||
TQKPSDLAYVIFTSGTTGKPKGVLEHQS VVQLRNSLIERYFGETNGSHAVLF LSNYVDFSLQLCLSVLG
2580      2590      2600      2610      2620      2630      2640      2650

```

X
R

```

GNKLIIPPEEGLTHEAFYDIGRREKLSYLSGTPSVLQQLIELSRLPHLMVT
X      2660      2670      2680      2690      2700

```

13. US-08-249-182-11 (1-23)
 P90430 Cyclophilin.

ID P90430 standard; peptide; 42 AA.
 AC P90430;
 DT 19-OCT-1989 (first entry)
 DE Cyclophilin.
 KW Cyclosporin A-binding; human cytosolic protein; chronic
 KW inflammatory disease; transplants; cyclophilin; rheumatoid arthritis;
 KW systemic lupus erythematosus; psoriasis; asthma; multiple sclerosis;
 KW myasthenia gravis; juvenile diabetes; autoimmune diseases.
 OS Homo sapiens (human).
 PN EP326067-A.
 PD 02-AUG-1989.
 PF 24-JAN-1989; 101164.
 PR 26-JAN-1988; US-148473.
 PA (DUPD) Du Pont De Nemours Co.
 PI Ackerman N R; Galbraith W; Irr J D; Jaffee B D; Lischwe M A.
 DR WPI; 89-222037/31.
 PT Cyclophilin and new isoforms
 PT - used as antiinflammatory and immuno-modulatory agent, eg to treat
 PT auto-immune diseases and prevent rejection.
 PS Claim 5; page 10; 12pp; English.
 CC Amino acids 1-42 of cyclosporin A-binding human cytosolic protein
 CC (17 kD, purity 95%). Val-1 is acetylated, or preceded by acetylated
 CC Met. This cyclophilin and its isoforms (see P90431-P90437) are used
 CC as anti-inflammatory and immuno-modulatory agents, eg to treat
 CC autoimmune diseases and to prevent rejection (see further uses
 CC in Keywords).
 SQ Sequence 42 AA;
 SQ 4 A; 2 R; 2 N; 4 D; 0 B; 0 C; 0 Q; 4 E; 0 Z; 2 G; 0 H;
 SQ 1 I; 3 L; 2 K; 0 M; 5 F; 3 P; 2 S; 4 T; 0 W; 0 Y; 4 V;

Initial Score = 7 Optimized Score = 8 Significance = 3.95
 Residue Identity = 44% Matches = 8 Mismatches = 10
 Gaps = 0 Conservative Substitutions = 0

```

      X 10      20 X
      TEFLSNYLTNVDDITLVPETLGR
      | ||| | |||
      VNPTVFFDITADDEPLGRVSFELFADKVPKTAENFRALSTGE
      X      10      20      30      40

```

14. US-08-249-182-11 (1-23)
 R32203 Apple fruit PPD pSR7.

ID R32203 standard; Protein; 68 AA.

AC R32203;
 DT 09-JUN-1993 (first entry)
 DE Apple fruit PPO pSR7.
 KW Polyphenol oxidase; PPO; catalyst; browning; fruit; plastid; vacuole;
 KW transform; coffee; tea; black olives; grapevine; chloroplast; apple;
 KW transit peptide; recombinant plasmid; PCR; broad bean; potato.
 OS Apple.
 PN W09302195-A.
 PD 04-FEB-1993.
 PF 16-JUL-1992; AV0356.
 PR 17-JUL-1991; AU-007248.
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PI Dry IB, Robinson SP;
 DR WPI; 93-058792/07.
 DR N-PSDB; Q36664.
 PT DNA encoding polyphenol oxidase polypeptide or fragment - useful
 PT for modifying the oxidase activity in fruit and vegetables to
 PT decrease or enhance browning
 PS Claim 8; Fig 3; 44pp; English.
 CC The sequences given in R32201-06 and R33772 represent polyphenol
 CC oxidase (PPO) enzymes from various plants. PPO is thought to be the
 CC predominant catalyst in browning of fruit caused by injury or damage.
 CC PPO is localised in the plastids of plant cells whereas the phenolic
 CC substrates of the enzyme are stored in the plant cell vacuole. This
 CC compartmentation prevents the browning reaction from occurring unless
 CC the plant cells are damaged and the enzyme and the substrate are
 CC mixed. The gene sequences encoding these proteins could be used to
 CC construct synthetic genes which may be used to transform plants to
 CC decrease expression of the enzyme gene. In some instances, eg.
 CC coffee, tea, black olives etc., it is desirable to increase the level
 CC of PPO to produce desired levels of browning or changes in flavour
 CC compounds. The grapevine PPO codes for an additional 103 amino acids
 CC upstream of the N-terminus of the mature protein. This region has the
 CC properties of a chloroplast transit peptide and is most likely
 CC responsible for targetting of the protein to be imported into the
 CC chloroplast and processed to produce mature PPO. Transformation of
 CC plants with this gene may therefore result in correct targetting and
 CC maturation of the grapevine PPO in other species and result in
 CC accumulation of active grapevine PPO enzyme in the plastids of these
 CC tissues.
 SQ Sequence 68 AA;
 SQ 4 A; 4 R; 5 N; 9 D; 0 B; 2 C; 0 Q; 4 E; 0 Z; 4 G; 2 H;
 SQ 0 I; 6 L; 4 K; 2 M; 4 F; 2 P; 3 S; 3 T; 3 W; 3 Y; 4 V;

Initial Score = 7 Optimized Score = 7 Significance = 3.95
 Residue Identity = 30% Matches = 7 Mismatches = 16
 Gaps = 0 Conservative Substitutions = 0

X 10 20 X
 TEFLSNYLTVDDITLVPETLGR
 ||| | |||
 EDMGNFYVSAGRDPLFYAHHCNVDRMWNVWKTLLGKKRKDPTDWDLDAEFLFYDENAELVSKVRDSL
 10 X 20 30 X 40 50 60

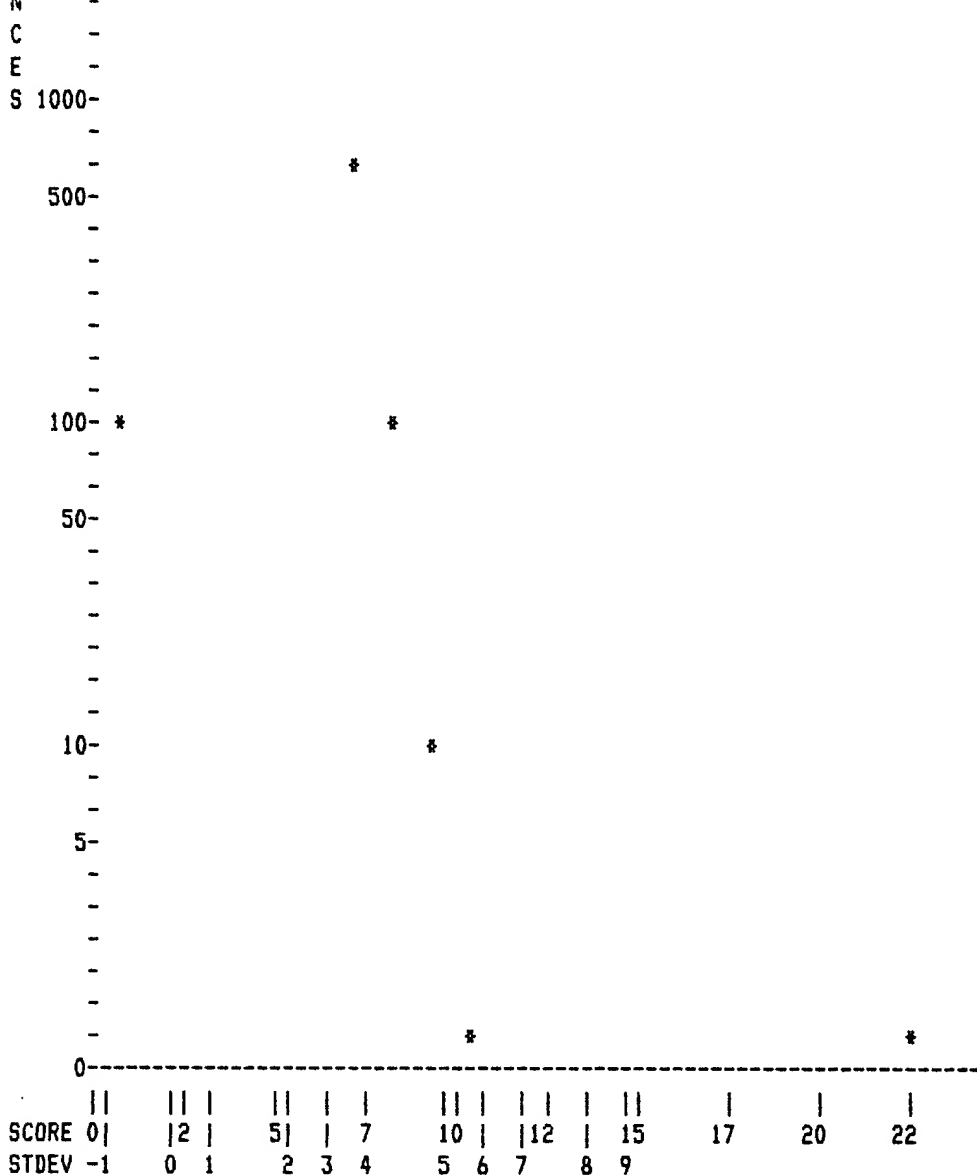
15. US-08-249-182-11 (1-23)
 R25119 Non-A, Non-B Hepatitis Virus antigen #10.

ID R25119 standard; Protein; 72 AA.
 AC R25119;
 DT 07-DEC-1992 (first entry)
 DE Non-A, Non-B Hepatitis Virus antigen #10.
 KW Antigen S29; NANBH; Hepatitis C; HCV; T064; T069; T06A; ELISA.
 OS Non-A Non-B Hepatitis Virus.
 PN W09209634-A.
 PD 11-JUN-1992.

```

100000-
      -
N       -
U50000-
M       -
B       -
E       *
R       -           *   *
      -
O       -
F10000-                   *
      -
S       -
E 5000-
Q       -
U       *
E       -

```



Cut-off raised to 5.
Cut-off raised to 6.
Cut-off raised to 7.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Init. Opt. Score	Sig.	Frame
**** 13 standard deviations above mean ****						
1. A42329	autotaxin - human (fragments)	114	22	22	14.00	0
**** 5 standard deviations above mean ****						
2. S33464	hypothetical protein - Arabid	341	10	10	5.16	0
**** 4 standard deviations above mean ****						
3. A44464	oxaloacetate decarboxylase be	141	9	10	4.42	0
4. S26196	imidazoleglycerol-phosphate d	208	9	9	4.42	0
5. DAPSPC	biphenyl-2,3-diol 1,2-dioxyge	293	9	9	4.42	0
6. C44465	sodium ion pump oxaloacetate	433	9	10	4.42	0
7. A27705	maltohexaose-producing amylas	518	9	9	4.42	0
8. S13163	65K antigen - Chinese hamster	556	9	9	4.42	0
9. 00BEHA	BSLF1 protein - saimiriine he	835	9	9	4.42	0
10. A35088	phycobiliprotein L-CM - Calot	1080	9	9	4.42	0
11. S11455	botulinum neurotoxin type D -	1276	9	9	4.42	0
**** 3 standard deviations above mean ****						
12. S20764	Ig heavy chain V region - hum	63	8	8	3.68	0
13. S04891	hypothetical protein, 9.2K -	82	8	8	3.68	0
14. S01741	hypothetical protein X1 - por	82	8	8	3.68	0
15. C60007	hypothetical protein C - porc	82	8	8	3.68	0
16. C36607	NS4 protein - porcine respira	82	8	8	3.68	0
17. S24280	hypothetical protein 4 - porc	82	8	8	3.68	0
18. S36265	Ig heavy chain V region (clon	118	8	8	3.68	0
19. S15611	hypothetical protein - sline	176	8	8	3.68	0
20. B25599	repB protein - Streptococcus	210	8	8	3.68	0
21. S05981	repB protein - Streptococcus	210	8	8	3.68	0
22. A40084	epidermal growth factor-relat	265	8	8	3.68	0
23. 00YV	transforming protein myb - av	265	8	8	3.68	0
24. S30811	sugar transport protein - yea	287	8	8	3.68	0
25. S36863	L-lactate dehydrogenase (EC 1	319	8	9	3.68	0
26. S35230	hypE protein - Bradyrhizobium	321	8	8	3.68	0
27. A43586	hypothetical protein 1 - Salm	345	8	9	3.68	0
28. S28129	gvpN protein - Halobacterium	347	8	8	3.68	0
29. S22387	cuticle-degrading proteinase	388	8	8	3.68	0
30. B48899	cephalosporinase, AmpC - Yers	388	8	8	3.68	0
31. S24562	beta-lactamase (EC 3.5.2.6) -	388	8	8	3.68	0
32. A36314	transforming protein (myb) -	388	8	8	3.68	0
33. A27891	RepA protein - Bacillus subti	396	8	8	3.68	0
34. PL0103	fibronectin receptor alpha ch	409	8	8	3.68	0
35. B25937	arsenical pump membrane prote	429	8	9	3.68	0
36. A47044	threonine dehydratase (EC 4.2	436	8	10	3.68	0
37. A47124	sporulation-specific 1,3-beta	445	8	8	3.68	0
38. A40639	exo-1,3-beta-glucanase - yeas	445	8	8	3.68	0
39. S36356	glucan 1,3-beta-glucosidase (445	8	8	3.68	0
40. JN0118	exo-1,3-beta-glucanase (EC 3.	448	8	8	3.68	0

1. US-08-249-182-11 (1-23)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
TITLE autotaxin - human (fragments)

ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
 Cioce, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis
 of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329
 ##status preliminary
 ##molecule_type protein
 ##residues 1-114 ##label STR
 ##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
 NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
 NCBIP:78509; NCBIP:78508; NCBIP:78503
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 114 #checksum 7335
 SEQUENCE

Initial Score = 22 Optimized Score = 22 Significance = 14.00
 Residue Identity = 95% Matches = 22 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

X      10      20 X
TEFLSNYLTNVDDITLVPETLGR
|||||
TEFLSNYLTNVDDITLVPGLGRDIEHLTSLDFFRVNSMGTVFVGYPGTFKGGQPLWITATKSPPFENINLY
X      10      20 X      30      40      50      60      70

```

Y

2. US-08-249-182-11 (1-23)

S33464 hypothetical protein - Arabidopsis thaliana

ENTRY S33464 #type complete
 TITLE hypothetical protein - Arabidopsis thaliana
 ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
 cress
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
 22-Nov-1993
 ACCESSIONS S33464
 REFERENCE S33464
 #authors Quigley, F.R.
 #submission submitted to the EMBL Data Library, May 1993
 #accession S33464
 ##status preliminary
 ##residues 1-341 ##label QUI
 ##cross-references EMBL:X71915
 SUMMARY #length 341 #molecular-weight 37008 #checksum 2677
 SEQUENCE

Initial Score = 10 Optimized Score = 10 Significance = 5.16
 Residue Identity = 43% Matches = 10 Mismatches = 13
 Gaps = 0 Conservative Substitutions = 0

```

X      10      20
TEFLSNYLTNVDDITLVPETLG
||| | ||| |||
FYDPASKKWTEVETTGAKPSARSVFAHAVVGKVIIFAGEVWPDNLNGHYGPGTSLNEGYALDTETLVWEKLG
230      240      250      260      270      280      290

```

X
R
EEGAPAI PRGWTAYTAATVDGKNGLLMHGGKLPNTERTDDLYFYAVNSA
X 300 310 320 330 340

3. US-08-249-182-11 (1-23)

A44464 oxaloacetate decarboxylase beta subunit (C terminu

ENTRY A44464 #type fragment
TITLE oxaloacetate decarboxylase beta subunit (C terminus) -
Klebsiella pneumoniae (fragment)
ORGANISM #formal_name Klebsiella pneumoniae
DATE 30-Apr-1993; #sequence_revision 30-Apr-1993; #text_change
30-Apr-1993
ACCESSIONS A44464
REFERENCE A44464
#authors Woehlke, G.; Laussermair, E.; Schwarz, E.; Desterhelt, D.;
Reinke, H.; Beyreuther, K.; Dimroth, P.
#journal J. Biol. Chem. (1992) 267:22804-22805
#title Appendix. Sequence of the beta-subunit of oxaloacetate
decarboxylase from Klebsiella pneumoniae: a correction of
the C-terminal part.
#cross-references MUID:93054592
#accession A44464
##status preliminary
##residues 1-141 ##label WOE
##cross-references NCBIP:118089
##note sequence extracted from NCBI backbone
SUMMARY #length 141 #checksum 9344
SEQUENCE

Initial Score = 9 Optimized Score = 10 Significance = 4.42
Residue Identity = 44% Matches = 11 Mismatches = 12
Gaps = 2 Conservative Substitutions = 0

X 10 20 X
TEF--LSNYLTNVDDITLVPETLGR
| | | | | | |
GNLMRESGVVERLSDTVQNALINIVTIFLGLSVGAKLVADKFLQPTLGILVLGVIAFCVGTAAAGVLMAKLM
10 20 X 30 40 50 60 70
NVFSRHKINPLIGSAGVSAVPMARVSN
80 90 100

4. US-08-249-182-11 (1-23)

S26196 imidazoleglycerol-phosphate dehydratase (EC 4.2.1.

ENTRY S26196 #type complete
TITLE imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) -
fungus (Trichoderma harzianum)
ORGANISM #formal_name Trichoderma harzianum
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
31-Dec-1993
ACCESSIONS S26196; S19787
REFERENCE S26196
#authors Goldman, G.H.; Demolder, J.; Dewaele, S.; Herrera-Estrella,
A.; Geremia, R.A.; van Montagu, M.; Contreras, R.
#journal Mol. Gen. Genet. (1992) 234:481-488
#title Molecular cloning of the imidazoleglycerolphosphate
dehydratase gene of Trichoderma harzianum by genetic
complementation in Saccharomyces cerevisiae using a direct
expression vector.
#accession S26196

##molecule_type mRNA
##residues 1-208 ##label GOL
##cross-references EMBL:Z11528

GENETICS

#gene igh
CLASSIFICATION #superfamily imidazoleglycerol-phosphate dehydratase
KEYWORDS carbon-oxygen lyase; histidine biosynthesis; hydro-lyase
SUMMARY #length 208 #molecular-weight 22356 #checksum 3322
SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTNVDDITLVPETLG
||| | | | |
KGDHLHIDDHTAEDCCIAVGTTFAKALGALTGVARFGYAYAPLDEALSRAVVDLSNRPTYVDLGLKREKLG
90 100 110 120 130 X 140 150

X
R

ELSCEMIPHCLQSFQAARITLHVDCLRGDNDHHRAESAFKALAVAVRWYD
160 170 180 190 200

5. US-08-249-182-11 (1-23)

DAPSPC biphenyl-2,3-diol 1,2-dioxygenase (EC 1.13.11.39)

ENTRY DAPSPC #type complete
TITLE biphenyl-2,3-diol 1,2-dioxygenase (EC 1.13.11.39) -
Pseudomonas sp.
ALTERNATE_NAMES 2,3-dihydroxybiphenyl dioxygenase
ORGANISM #formal_name Pseudomonas sp.
DATE 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change
30-Jun-1993
ACCESSIONS A32312; JU0085
REFERENCE A32312
#authors Kimbara, K.; Hashimoto, T.; Fukuda, M.; Koana, T.; Takagi,
M.; Oishi, M.; Yano, K.
#journal J. Bacteriol. (1989) 171:2740-2747
#title Cloning and sequencing of two tandem genes involved in
degradation of 2,3-dihydroxybiphenyl to benzoic acid in the
polychlorinated biphenyl-degrading soil bacterium
Pseudomonas sp. strain KKS102.
#cross-references MUID:89213965
#contents Strain KKS102
#accession A32312
##molecule_type DNA
##residues 1-293 ##label KIM
REFERENCE JU0085
#authors Takagi, M.
#submission submitted to JIPID, July 1989
#accession JU0085
##molecule_type DNA
##residues 1-293 ##label TAK
##note the submitted nucleotide sequence was translated at
JIPID

COMMENT This enzyme catalyzes the third step in the major degradative
pathway for biphenyl and polychlorinated biphenyls (PCBs),
cleavage of a 2,3-dihydroxybiphenyl derivative at the 1 and 2
positions to give a derivative of 2-hydroxy-6-oxo-6-phenylhexa-2,
4-dienoate.

GENETICS

#gene bphC

CLASSIFICATION #superfamily biphenyl-2,3-diol 1,2-dioxygenase
 KEYWORDS oxidoreductase; PCB biodegradation
 SUMMARY #length 293 #molecular-weight 32244 #checksum 7776
 SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYLTVDDITLVPETLG
                                ||| | || |||
LPLEIYYGPAEIFHEPFLPSAPVSGFVTGDGGIGHFVRCVPDTAKAMAFYTEVLGFVLSIDIIDIMGPETSV
  120      130      140      150      160 X      170      180

X
R

PAHFLHCNGRHHHTIALAAFFIPKRIHHFMLQANTIDDVGYAFDRLDAAGRI
X  190      200      210      220      230

```

6. US-08-249-182-11 (1-23)

C44465 sodium ion pump oxaloacetate decarboxylase subunit

ENTRY C44465 #type complete
 TITLE sodium ion pump oxaloacetate decarboxylase subunit beta -
 Salmonella typhimurium
 ORGANISM #formal_name Salmonella typhimurium
 DATE 30-Apr-1993; #sequence_revision 30-Apr-1993; #text_change
 30-Apr-1993
 ACCESSIONS C44465
 REFERENCE A44465
 #authors Woehlke, G.; Wifling, K.; Dimroth, P.
 #journal J. Biol. Chem. (1992) 267:22798-22803
 #title Sequence of the sodium ion pump oxaloacetate decarboxylase
 from Salmonella typhimurium.
 #cross-references MUID:93054591
 #contents LT2
 #accession C44465
 ##status preliminary
 ##residues 1-433 ##label WOE
 ##cross-references NCBIP:118073
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 433 #molecular-weight 44923 #checksum 4940
 SEQUENCE

Initial Score = 9 Optimized Score = 10 Significance = 4.42
 Residue Identity = 44% Matches = 11 Mismatches = 12
 Gaps = 2 Conservative Substitutions = 0

```

                                X      10      20
                                TEF--LSNYLTNVDDITLVPET
                                || || || || ||
FPVVLLMLVALLLPDAAPLLGMFCFGLNMRSGVVERLSDTVQNGLINIVTIFLGLSVGAKLVADKFLQPGT
  270      280      290      300      310      320      330

X
LGR
||
LGILLLGVIAGIGTAAGVLMKLLNLCSKNKINPLIGSAGVSAVPMARVSN
340 X      350      360      370      380      390

```

7. US-08-249-182-11 (1-23)

A27705 maltotetraose-producing amylase (EC 3.2.1.-) precursor

ENTRY A27705 #type complete
 TITLE maltohexaose-producing amylase (EC 3.2.1.-) precursor -
 Bacillus sp.
 ALTERNATE_NAMES G6-amylase
 ORGANISM #formal_name Bacillus sp.
 DATE 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change
 28-May-1993
 ACCESSIONS A27705
 REFERENCE A27705
 #authors Tsukamoto, A.; Kimura, K.; Ishii, Y.; Takano, T.; Yamane, K.
 #journal Biochem. Biophys. Res. Commun. (1988) 151:25-31
 #title Nucleotide sequence of the maltohexaose-producing amylase
 gene from an alkalophilic Bacillus sp. #707 and structural
 similarity to liquefying type alpha-amylases.
 #cross-references MUID:88162814
 #contents Strain #707
 #accession A27705
 ##molecule_type DNA
 ##residues 1-518 ##label TSU
 KEYWORDS glycosidase; hydrolase
 FEATURE
 1-33 #domain signal sequence #label SIG\
 34-518 #protein maltohexaose-producing amylase #label MAT
 SUMMARY #length 518 #molecular-weight 59008 #checksum 7204
 SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

```

      X      10      20 X
      TEFLSNVLTNVDDITLVPETLGR
          ||| | || |||
MKHRTGKKGFLSILLAFLLVITSIPFTLVDVEAHHNGTNGTMMQYFEWYLPNDGNHWNRLNSDASNLKSKGI
      10      20      30      40      50      60      70

TAVWIPPA
      80

```

8. US-08-249-182-11 (1-23)
 S13163 65K antigen - Chinese hamster

ENTRY S13163 #type complete
 TITLE 65K antigen - Chinese hamster
 ORGANISM #formal_name Cricetulus griseus #common_name Chinese hamster
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change
 21-Nov-1993
 ACCESSIONS S13163
 REFERENCE S13163
 #authors Ahmad, S.; Gupta, R.S.
 #journal Biochim. Biophys. Acta (1990) 1087:253-255
 #title Cloning of a Chinese hamster protein homologous to the mouse
 t-complex protein TCP-1: structural similarity to the
 ubiquitous 'chaperonin' family of heat-shock proteins.
 #cross-references MUID:91027940
 #accession S13163
 ##status preliminary
 ##residues 1-556 ##label AHM
 SUMMARY #length 556 #molecular-weight 60338 #checksum 2691
 SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
|| | | |
RAFHNEAQVNPENKLNKWLGLDLINGKPRDNKQAGVFPTIVKVKSLKFATEAAITILRIDDLIKLHPETKD
470 480 490 500 510 520 530

X
R

DKHGSYENAVHSGALDD
540 550

9. US-08-249-182-11 (1-23)

00BEHA BSLF1 protein - saimiriine herpesvirus 1 (strain 1)

ENTRY 00BEHA #type complete
TITLE BSLF1 protein - saimiriine herpesvirus 1 (strain 11)
ORGANISM #formal_name saimiriine herpesvirus 1
 #note host Saimiri sciureus (common squirrel monkey)
DATE 30-Sep-1989 #sequence_revision 31-Dec-1992 #text_change
 04-Mar-1994
ACCESSIONS I36811
REFERENCE A36806
 #authors Albrecht, J.
 #submission submitted to the EMBL Data Library, January 1992
 #description Primary structure of the herpesvirus saimiri genome.
 #accession I36811
 ##molecule_type DNA
 ##residues 1-835 ##label ALB
 ##cross-references GB:X64346
REFERENCE A37309
 #authors Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.;
 Biesinger, B.; Newman, C.; Wittmann, S.; Craxton, M.A.;
 Coleman, H.; Fleckenstein, B.; Honess, R.W.
 #journal J. Virol. (1992) 66:5047-5058
 #title Primary structure of the herpesvirus saimiri genome.
 #cross-references MUID:92333688
 #contents annotation; possible protein-coding frames
 #note neither protein nor nucleotide sequence is given in this
 paper
GENETICS
 #gene 56
CLASSIFICATION #superfamily varicella-zoster virus gene 6 protein
SUMMARY #length 835 #molecular-weight 96127 #checksum 1688
SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
||| | | |
NYCHIKLARDSLESQAIDTSIDTLRGQLMSNQLVHYIYLSFFQCLNKDIFIYKSHLTNSDNIHFVPETEV
240 250 260 270 280 X 290 300

X
R

LAGSLDENFRKDMITYYKSTYLKTYITHKCIHLPDLIGYAPQDCTSFVYH
X 310 320 330 340 350

10. US-08-249-182-11 (1-23)

A35088 phycobiliprotein L-CM - Calothrix sp.

ENTRY A35088 #type complete
TITLE phycobiliprotein L-CM - Calothrix sp.
ORGANISM #formal_name Calothrix sp.
DATE 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change
18-Jun-1993
ACCESSIONS A35088
REFERENCE A35088
#authors Houmard, J.; Capuano, V.; Colombano, M.V.; Coursin, T.;
Tandeau de Marsac, N.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:2152-2156
#title Molecular characterization of the terminal energy acceptor of
cyanobacterial phycobilisomes.
#cross-references MUID:90192765
#accession A35088
##status preliminary
##molecule_type DNA
##residues 1-1080 ##label HOU
##cross-references GB:M20806; GB:M31224
SUMMARY #length 1080 #molecular-weight 120456 #checksum 6806
SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNVLTNVDDITLVPETLG
|| || || | ||
IRKYNQILATGGIRAFIGALVSSAEYAEVFGEDTVPYRRYPTLPAANFPNTEKLYNQLTQNDLVPSPFKT
830 840 850 860 870 X 880 890

X
R

VQPRLTLAGTSSSGRNGFTDLGRSSTSAQGQLGETANRCKPARIYRLSGTN
900 910 920 930 940

11. US-08-249-182-11 (1-23)

S11455 botulinum neurotoxin type D - Clostridium botulinu

ENTRY S11455 #type complete
TITLE botulinum neurotoxin type D - Clostridium botulinum
ORGANISM #formal_name Clostridium botulinum
DATE 18-Feb-1994; #sequence_revision 18-Feb-1994; #text_change
18-Feb-1994
ACCESSIONS S11455
REFERENCE S11455
#authors Binz, T.; Kurazono, H.; Popoff, M.R.; Eklund, M.W.;
Sakaguchi, G.; Kozaki, S.; Krieglstein, K.; Henschen, A.;
Gill, D.M.; Niemann, H.
#journal Nucleic Acids Res. (1990) 18:5556
#title Nucleotide sequence of the gene encoding Clostridium
botulinum neurotoxin type D.
#cross-references MUID:91016853
#accession S11455
##status preliminary
##residues 1-1276 ##label BIN
##cross-references EMBL:X54254
SUMMARY #length 1276 #molecular-weight 146871 #checksum 326
SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 4.42
Residue Identity = 39% Matches = 9 Mismatches = 14

Gaps = 0 Conservative Substitutions = 0

```

X      10      20
TEFLSNYLTNVDDITLVPETLG
||| | |||
DGQVPINPEIVDPLLPNVNMEPLNLPGEIIVFYDDITKYVDYLNQVYVLESQKLSNNVENITLTTSVEEALG
500      510      520      530      540      550      560

X
R

YSNKIYTFPLSLAEKVNKGVQAGLFLNWANEVVEDFTTNIMKKDTLDKISD
570      580      590      600      610      620

```

12. US-08-249-182-11 (1-23)

S20764 Ig heavy chain V region - human

```

ENTRY      S20764      #type complete
TITLE      Ig heavy chain V region - human
ORGANISM    #formal_name Homo sapiens #common_name man
DATE        19-Feb-1994; #sequence_revision 19-Feb-1994; #text_change
            19-Feb-1994
ACCESSIONS  S20764
REFERENCE   S20764
#authors    Mortari, F.; Wang, J.; Schroeder, H.W.
#submission submitted to the EMBL Data Library, April 1992
#accession  S20764
##status    preliminary
##residues  1-63 ##label MOR
##cross-references EMBL:Z11951
SUMMARY     #length 63 #molecular-weight 6858 #checksum 190
SEQUENCE

```

Initial Score = 8 Optimized Score = 8 Significance = 3.68
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

```

X      10      20 X
TEFLSNYLTNVDDITLVPETLGR
|| | ||| ||
RKSDGGTTDYAAPVKGRFTISRDDSKNTVYLQMNLSKLTEDSGVYVCTDDIGWGQGLVTVSS
10      20      30      40      50      60

```

13. US-08-249-182-11 (1-23)

S04891 hypothetical protein, 9.2K - porcine transmissible

```

ENTRY      S04891      #type complete
TITLE      hypothetical protein, 9.2K - porcine transmissible
            gastroenteritis virus
ORGANISM    #formal_name porcine transmissible gastroenteritis virus
DATE        30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
            18-Jun-1993
ACCESSIONS  S04891
REFERENCE   S04889
#authors    Kapke, P.A.; Tung, F.Y.T.; Brian, D.A.
#submission submitted to the EMBL Data Library, September 1988
#description Nucleotide sequence between the peplomer and matrix protein
            genes of the porcine transmissible gastroenteritis
            coronavirus identifies three large open reading frames.
#accession  S04891
##molecule_type DNA
##residues  1-82 ##label KAP
##cross-references EMBL:X12800
SUMMARY     #length 82 #molecular-weight 9239 #checksum 9602

```

SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 3.68
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

      X      10      20  X
      TEFLSNYLTNVDDITLVPETLGR
          || || | |||
MTFPRALTVIDDNGMVINIIFWFLIIILILLSIALLNIIKLCMVCCNLGRTVIIIVPAQHAYDAYKNFMRIK
      10      20      30      40      50      60      70

AYNPDGALLA
      80
```

14. US-08-249-182-11 (1-23)

S01741 hypothetical protein X1 - porcine transmissible

ENTRY S01741 #type complete
TITLE hypothetical protein X1 - porcine transmissible
gastroenteritis virus (strain Purdue-115)
ORGANISM #formal_name porcine transmissible gastroenteritis virus
DATE 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change
18-Jun-1993
ACCESSIONS S01741
REFERENCE S01738
#authors Rasschaert, D.; Gelfi, J.; Laude, H.
#journal Biochimie (1987) 69:591-600
#title Enteric coronavirus TGEV: partial sequence of the genomic
RNA, its organization and expression.
#cross-references MUID:88078100
#accession S01741
##molecule_type mRNA
##residues 1-82 ##label RAS
##cross-references EMBL:X06371
SUMMARY #length 82 #molecular-weight 9239 #checksum 9602
SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 3.68
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

      X      10      20  X
      TEFLSNYLTNVDDITLVPETLGR
          || || | |||
MTFPRALTVIDDNGMVINIIFWFLIIILILLSIALLNIIKLCMVCCNLGRTVIIIVPAQHAYDAYKNFMRIK
      10      20      30      40      50      60      70

AYNPDGALLA
      80
```

15. US-08-249-182-11 (1-23)

C60007 hypothetical protein C - porcine transmissible

ENTRY C60007 #type complete
TITLE hypothetical protein C - porcine transmissible
gastroenteritis virus (strain virulent Miller)
ORGANISM #formal_name porcine transmissible gastroenteritis virus
DATE 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
30-Sep-1993
ACCESSIONS C60007
REFERENCE A60007
#authors Wesley, R.D.; Cheung, A.K.; Michael, D.D.; Woods, R.D.
#journal Virus Res. (1989) 13:87-100

#title Nucleotide sequence of coronavirus TGEV genomic RNA: evidence
for 3 mRNA species between the peplomer and matrix protein
genes.

#accession C60007

##molecule_type mRNA

##residues 1-82 ##label WES

SUMMARY #length 82 #molecular-weight 9268 #checksum 238

SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 3.68
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

X 10 20 X
TEFLSNYLTNVDDITLVPETLGR
|| || | ||
MTFPRALTVIDDNGMVISIIFWFLIIILILLSIALLNIIKLCMVCCNLGRTVIIIPVQHAYDAYKNFMRIK
10 20 30 40 50 60 70

AYNPDGALLV
80

> 0 <
0| |0 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_11s.res made by on Thu 22 Sep 94 10:14:07-PDT.

Query sequence being compared:US-08-249-182-11 (1-23)
Number of sequences searched: 36000
Number of scores above cutoff: 4784

Results of the initial comparison of US-08-249-182-11 (1-23) with:
Data bank : Swiss-Prot 28, all entries

10000- * * *
-
N - *
U 5000-
M -
B -
E -
R -
- *
O -
F 1000-
-
S *
E 500-
Q -
U - *
E -
N -
C -
E -
S 100-
-
-
50- *
-
-
- *
-

10-

*

5-

0-

SCORE	0	1	2	3	4	5	6	7	8	9
STDEV		-1	0	1	2		3	4		

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	3	4	1.28

Times:	CPU	Total Elapsed
	00:00:53.89	00:01:00.00

Number of residues:	12496420
Number of sequences searched:	36000
Number of scores above cutoff:	4784

Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.
 Cut-off raised to 6.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
1. HIS7_TRIHA	IMIDAZOLEGLYCEROL-PHOSPHATE D	208	9	9	4.69	0
2. BPHC_PSES1	BIPHENYL-2,3-DIOL 1,2-DIOXYGE	293	9	9	4.69	0
3. DCOB_SALTY	OXALOACETATE DECARBOXYLASE BE	433	9	10	4.69	0
4. AMT6_BACS7	GLUCAN 1,4-ALPHA-MALTOHEXAOSI	518	9	9	4.69	0
5. TCP1_CRIGR	T-COMPLEX PROTEIN 1 (TCP-1) (556	9	9	4.69	0
6. UL52_HSVSA	PROBABLE DNA REPLICATION GENE	835	9	9	4.69	0
7. APCE_SYNY4	PHYCOBILISOME 120 KD LINKER P	896	9	9	4.69	0

8. APCE_FREDI	PHYCOBILISOME 120 KD LINKER P	1079	9	9	4.69	0
9. BXD_CLOBO	BOTULINUM NEUROTOXIN TYPE D P	1276	9	9	4.69	0
**** 3 standard deviations above mean ****						
10. VNS4_CVPRM	NONSTRUCTURAL PROTEIN 4 (X1 P	82	8	8	3.91	0
11. VNS4_CVPPU	NONSTRUCTURAL PROTEIN 4 (X1 P	82	8	8	3.91	0
12. REPB_STRPN	REPLICATION PROTEIN REPB.	210	8	8	3.91	0
13. HYPE_BRAJA	HYPE PROTEIN.	321	8	8	3.91	0
14. EGF2_STRPU	EPIDERMAL GROWTH FACTOR-RELAT	325	8	8	3.91	0
15. MYB_AVIMB	MYB TRANSFORMING PROTEIN.	382	8	8	3.91	0
16. CUDP_METAN	CUTICLE-DEGRADING PROTEASE PR	388	8	8	3.91	0
17. REPA_BACSU	REPA PROTEIN.	396	8	8	3.91	0
18. ITA5_MOUSE	FIBRONECTIN RECEPTOR ALPHA SU	409	8	8	3.91	0
19. ARSB_ECOLI	ARSENICAL PUMP MEMBRANE PROTE	429	8	9	3.91	0
20. THD1_CORGL	THREONINE DEHYDRATASE BIOSYNT	436	8	10	3.91	0
21. SPR1_YEAST	SPORULATION-SPECIFIC GLUCAN 1	445	8	8	3.91	0
22. EXG1_YEAST	GLUCAN 1,3-BETA-GLUCOSIDASE I	448	8	8	3.91	0
23. MKS1_YEAST	NEGATIVE REGULATOR OF RAS-CAM	458	8	8	3.91	0
24. US15_HCMVA	HYPOTHETICAL PROTEIN HVLFG.	484	8	8	3.91	0
25. TRPE_PSEAE	ANTHRANILATE SYNTHASE COMPONE	492	8	8	3.91	0
26. TRPE_PSEPU	ANTHRANILATE SYNTHASE COMPONE	493	8	8	3.91	0
27. AMIC_STRPN	OLIGOPEPTIDE TRANSPORT AMIC P	498	8	9	3.91	0
28. SYFD_YEAST	PHENYLALANYL-TRNA SYNTHETASE	502	8	8	3.91	0
29. TRPE_PSESS	ANTHRANILATE SYNTHASE COMPONE	505	8	8	3.91	0
30. ACHD_RAT	ACETYLCHOLINE RECEPTOR, DELTA	517	8	8	3.91	0
31. ACHD_MOUSE	ACETYLCHOLINE RECEPTOR PROTEI	520	8	8	3.91	0
32. VNS1_AHSV4	NONSTRUCTURAL PROTEIN NS1 (HY	548	8	8	3.91	0
33. TCPB_MOUSE	T-COMPLEX PROTEIN 1 (TAILLESS	556	8	8	3.91	0
34. TCP1_RAT	T-COMPLEX PROTEIN 1 (TCP-1).	556	8	8	3.91	0
35. TCP1_HUMAN	T-COMPLEX PROTEIN 1 (TCP-1).	556	8	8	3.91	0
36. TCP1_DROME	T-COMPLEX PROTEIN 1 HOMOLOG (557	8	8	3.91	0
37. MYB_CHICK	MYB PROTO-ONCOGENE PROTEIN.	641	8	8	3.91	0
38. VID4_AGRTU	VIRD4 PROTEIN PRECURSOR.	656	8	8	3.91	0
39. VID4_AGRT5	VIRD4 PROTEIN.	665	8	8	3.91	0
40. VID4_AGRRA	VIRD4 PROTEIN PRECURSOR.	671	8	8	3.91	0

1. US-08-249-182-11 (1-23)

HIS7_TRIHA IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRATASE (EC 4.2.1.

ID HIS7_TRIHA STANDARD; PRT; 208 AA.
AC P34041;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRATASE (EC 4.2.1.19).
GN IGH.
OS TRICHODERMA HARZIANUM.
OC EUKARYOTA; FUNGI; DEUTEROMYCOTINA (IMPERFECT FUNGI).
RN [1]
RP SEQUENCE FROM N.A.
RM 93024323
RA GOLDMAN G.H., DEMOLDER J., DEWAELE S., HERRERA-ESTRELLA A.,
RA GEREMIA R.A., VAN MONTAGU M., CONTRERAS R.;
RL MOL. GEN. GENET. 234:481-488(1992).
CC -!- CATALYTIC ACTIVITY: D-ERYTHRO-1-(IMIDAZOL-4-YL)GLYCEROL
CC 3-PHOSPHATE = 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + H(2)O.
CC -!- PATHWAY: SEVENTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: TO OTHER IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRATASES.
DR EMBL; Z11528; THIGPMR.
DR PIR; S26196; S26196.
KW HISTIDINE BIOSYNTHESIS; LYASE.
SQ SEQUENCE 208 AA; 22356 MW; 196910 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
||| | | | |
KGDHLHIDDHTAEDCCIAVGTTFAKALGALTGVARFGVAYAPLDEALSRVVDLSNRPTYVVDLGLKREKLG
90 100 110 120 130 X 140 150

X
R

ELSCEMIPHCLQSFQAARITLHVDCLRGDNDHHRAESAFKALAVAVRWYD
160 170 180 190 200

2. US-08-249-182-11 (1-23)

BPHC_PSES1 BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.39)

ID BPHC_PSES1 STANDARD; PRT; 293 AA.
AC P17297;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.39) (23DHBP OXYGENASE)
DE (2,3-DIHYDROXYBIPHENYL DIOXYGENASE).
GN BPHC.
OS PSEUDOMONAS SP. (STRAIN KKS102).
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC PSEUDOMONADACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 89213965
RA KIMBARA K., HASHIMOTO T., FUKUDA M., KOANA T., TAKAGI M., OISHI M.,
RA YANO K.;
RL J. BACTERIOL. 171:2740-2747(1989).
CC -!- CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-
CC 6-PHENYLHEXA-2,4-DIENOATE + H(2)O.
CC -!- PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO
CC BENZOIC ACID AND CHLOROBENZOIC ACIDS.
CC -!- SUBUNIT: HOMOOCTAMER.
CC -!- COFACTOR: REQUIRES FERROUS IRON AS THE PROSTHETIC GROUP.
CC -!- SIMILARITY: WITH OTHER EXTRADIOL RING-CLEAVAGE DIOXYGENASES.
DR EMBL; M26433; M26433.
DR PIR; A32312; DAPSPC.
DR PROSITE; PS00082; EXTRADIOL_DIOXYGENAS.
KW OXIDOREDUCTASE; DIOXYGENASE; AROMATIC HYDROCARBONS CATABOLISM; IRON.
SQ SEQUENCE 293 AA; 32244 MW; 433312 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
||| | | | |
LPLEIYYGPAEIFHEPFLPSAPVSGFVTGDGGIGHFVRCVPDTAKAMAFYTEVLGFVLSDIIDIQMGPESTV
120 130 140 150 160 X 170 180

X
R

PAHFLHCNCRHHTIALAAFFIPKRIHHFMLQANTIDVGYAFDRDLAAGRI
X 190 200 210 220 230

3. US-08-249-182-11 (1-23)

DCOB_SALTY OXALOACETATE DECARBOXYLASE BETA CHAIN (EC 4.1.1.3)

ID DCOB_SALTY STANDARD; PRT; 433 AA.
AC 003031;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE OXALOACETATE DECARBOXYLASE BETA CHAIN (EC 4.1.1.3).
GN OADB.
OS SALMONELLA TYPHIMURIUM.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 93054591
RA WOHLKE G., WIFLING K., DIMROTH P.;
RL J. BIOL. CHEM. 267:22798-22803(1992).
CC -!- FUNCTION: LYASE AND SODIUM TRANSPORTER.
CC -!- CATALYTIC ACTIVITY: OXALOACETATE = PYRUVATE + CO(2).
CC -!- SUBUNIT: COMPOSED OF THREE CHAINS (ALPHA, BETA, AND GAMMA).
CC -!- COFACTOR: REQUIRES A SODIUM ION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE-BOUND.
DR EMBL; M96434; ST0ADGABA.
DR PIR; C44465; C44465.
KW DECARBOXYLASE; TRANSMEMBRANE; LYASE; SODIUM TRANSPORT.
FT TRANSMEM 22 45 I (BY SIMILARITY).
FT TRANSMEM 121 145 II (BY SIMILARITY).
FT TRANSMEM 159 186 III (6-HELIX MODEL) (BY SIMILARITY).
FT TRANSMEM 159 180 IIIA (7-HELIX MODEL) (BY SIMILARITY).
FT TRANSMEM 193 213 IIIB (7-HELIX MODEL) (BY SIMILARITY).
FT TRANSMEM 214 241 IV (6-HELIX MODEL) (BY SIMILARITY).
FT TRANSMEM 218 241 IV (7-HELIX MODEL) (BY SIMILARITY).
FT TRANSMEM 266 293 V (BY SIMILARITY).
SQ SEQUENCE 433 AA; 44923 MW; 956333 CN;

Initial Score = 9 Optimized Score = 10 Significance = 4.69
Residue Identity = 44% Matches = 11 Mismatches = 12
Gaps = 2 Conservative Substitutions = 0

					X	10	20
					TEF--LSNYLTNVDDITLVPET		
FPVVLLMLVALLLPDAAPLLGMFCFGNLMRESGVVERLSDTVQNGLINIVTIFLGLSVGAKLVADKFLQPT							
270	280	290	300	310	320	330	

					X
					LGR
LGILLGVIAGGIGTAAGVLMAKLLNLCSKNKINPLIGSAGVSAPVMAARVSN					
340 X	350	360	370	380	390

4. US-08-249-182-11 (1-23)

AMT6_BACS7 GLUCAN 1,4-ALPHA-MALTOHEXAOSIDASE PRECURSOR (EC 3.

ID AMT6_BACS7 STANDARD; PRT; 518 AA.
AC P19571;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE GLUCAN 1,4-ALPHA-MALTOHEXAOSIDASE PRECURSOR (EC 3.2.1.98) (G6-AMYLASE)
DE (MALTOHEXAOSIDE-PRODUCING AMYLASE) (EXO-MALTOHEXAOSIDASE).
OS BACILLUS SP. (STRAIN 707).
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-36.
RM 88162814

RA YSUKARUO A.; KIMURA K.; ISHII Y.; TAKANO T.; YAMANE K.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 151:25-31(1988).
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC LINKAGES
 CC IN AMYLACEOUS POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE
 CC MALTOHEXAOSE RESIDUES FROM THE NON-REDUCING CHAIN ENDS.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- PATHWAY: DEGRADATION OF STARCH.
 CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL; M18862; BSANYG6.
 DR PIR; A27705; A27705.
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.
 FT SIGNAL 1 33
 FT CHAIN 34 518 MALTOHEXAOSE-PRODUCING AMYLASE.
 SQ SEQUENCE 518 AA; 59009 MW; 1397695 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

X 10 20 X
 TEFLSNVLTNVDDITLVPETLGR
 ||| | || ||
 MKMRTGKKGFLSILLAFLLVITSIPFTLV DVEAHHNGTNGTMMQYFEWYLPNDGNHWNRLNSDASNLKSKGI
 10 20 30 40 50 60 70
 TAVWIPPA
 80

5. US-08-249-182-11 (1-23)

TCP1_CRIGR T-COMPLEX PROTEIN 1 (TCP-1) (65 KD ANTIGEN).

ID TCP1_CRIGR STANDARD; PRT; 556 AA.
 AC P18279;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE T-COMPLEX PROTEIN 1 (TCP-1) (65 KD ANTIGEN).
 OS CRICETULUS GRISEUS (CHINESE HAMSTER).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91027940
 RA AHMAD S., GUPTA R.S.;
 RL BIOCHIM. BIOPHYS. ACTA 1087:253-255(1990).
 CC -!- FUNCTION: MOLECULAR CHAPERONE. KNOWN TO PLAY A ROLE, IN VITRO,
 CC IN THE FOLDING OF ACTIN AND TUBULIN.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SUBUNIT: FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER;
 CC ASSOCIATED WITH OTHER PROTEINS IN 850 KD TO 900 KD COMPLEX.
 CC -!- SIMILARITY: TO OTHER MEMBERS OF TCP-1 CHAPERONIN FAMILY.
 DR EMBL; M34665; CGTCP1A.
 DR PIR; S13163; S13163.
 DR PROSITE; PS00750; TCP1_1.
 DR PROSITE; PS00751; TCP1_2.
 KW CHAPERONE.
 SQ SEQUENCE 556 AA; 60338 MW; 1473942 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNVLTNVDDITLVPETLG

RAFHNEAGVNPGRKNLKWIGLDLNGKPRDNKQAGVFEPTIVKVKSLKFATEAAITILRIDDLIKLHPETKD
470 480 490 500 510 520 530

X
R

DKHGSYENAVHSGALDD
540 550

6. US-08-249-182-11 (1-23)

UL52_HSVSA PROBABLE DNA REPLICATION GENE 56 PROTEIN.

ID UL52_HSVSA STANDARD; PRT; 835 AA.
AC P14346;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE PROBABLE DNA REPLICATION GENE 56 PROTEIN.
GN 56 OR EDRF4.
OS HERPESVIRUS SAIMIRI (STRAIN 11).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 92333688
RA ALBRECHT J.-C., NICHOLAS J., BILLER D., CAMERON K.R., BIESINGER B.,
RA NEWMAN C., WITTMANN S., CRAXTON M.A., COLEMAN H., FLECKENSTEIN B.,
RA HONESS R.W.;
RL J. VIROL. 66:5047-5058(1992).
RN [2]
RP SEQUENCE FROM N.A.
RM 92230228
RA NICHOLAS J., CAMERON K.R., COLEMAN H., NEWMAN C., HONESS R.W.;
RL VIROLOGY 188:296-310(1992).
RN [3]
RP SEQUENCE OF 631-835 FROM N.A.
RM 88300875
RA NICHOLAS J., GOMPELS U.A., CRAXTON M.A., HONESS R.W.;
RL J. VIROL. 62:3250-3257(1988).
CC -!- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL52,
CC EBV-1 7, EBV BSLF1, HVS-1 56, HCMV UL70 AND VZV 6.
DR EMBL; X64346; HSGEND.
DR EMBL; M86409; HEHSV3PRG.
DR EMBL; M21943; HEHSS52K.
DR PIR; I36811; Q0BEHA.
KW DNA REPLICATION.
SQ SEQUENCE 835 AA; 96127 MW; 3760278 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
Residue Identity = 39% Matches = 9 Mismatches = 14
Gaps = 0 Conservative Substitutions = 0

X 10 20
TEFLSNYLTVDDITLVPETLG
||| || ||||
NYCHIKLARDSLESQAIDTSIDTLRGQLMSNODLVHYIYLSFFQCLNKDIFIKYSHLTNSDNIHFVPETEV
240 250 260 270 280 X 290 300

X
R

LAQSLDENFRKDMITYYKSTYLKTYITHKCIHLPDLIGYAPQDCTSFVYH
X 310 320 330 340 350

7. US-08-249-182-11 (1-23)

APCE SYNY4 PHYCOBILISOME 120 KD LINKER POLYPEPTIDE, CORE (L-C

```

ID      APCE_SYNV4      STANDARD;      PRT;      896 AA.
AC      002907;
DT      01-JUL-1993 (REL. 26, CREATED)
DT      01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT      01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE      PHYCOBILISOME 120 KD LINKER POLYPEPTIDE, CORE (L-CM 92) (CORE-MEMBRANE
DE      LINKER PROTEIN).
OS      SYNECHOCYSTIS SP. (STRAIN PCC 6714).
OC      PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC      CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.
RN      [1]
RP      SEQUENCE FROM N.A.
RM      93222481
RA      DIMAGNO L.M., HASELKORN R.;
RL      PLANT MOL. BIOL. 21:835-846(1993).
CC      -!- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL
CC      ENERGY ACCEPTOR (BY ITS PHYCOBILIN-LIKE DOMAINS) AND AS A LINKER
CC      POLYPEPTIDE (BY ITS REPEATS AND ARMS) THAT STABILIZES THE
CC      PHYCOBILISOME CORE ARCHITECTURE.
CC      -!- SUBUNIT: PHYCOBILISOMES OF THIS ORGANISM ARE COMPOSED OF A TWO
CC      CYLINDER CORE, FROM WHICH SIX RODS RADIATE. THE CORE IS MAINLY
CC      COMPOSED OF ALLOPHYCOCYANIN ALPHA AND BETA CHAINS AND OF MINOR
CC      COMPONENTS.
CC      -!- SUBCELLULAR LOCATION: ANCHORS THE PHYCOBILISOME PERPENDICULARLY
CC      TO THE STROMAL SURFACE OF THE THYLAKOID MEMBRANE.
CC      -!- SIMILARITY: THE REPEATED DOMAINS ARE SIMILAR TO THE N-TERMINAL
CC      REGIONS OF PHYCOCYANIN ROD LINKER POLYPEPTIDES.
CC      -!- SIMILARITY: THE PHYCOBILIN-LIKE DOMAINS ARE SIMILAR TO PHYCOBILINS
CC      FROM VARIOUS SPECIES.
DR      EMBL; L02309; SSLCM.
KW      PHYCOBILISOME; ELECTRON TRANSPORT; PHOTOSYNTHESIS; REPEAT.
SQ      SEQUENCE      896 AA;  100460 MW;  4106962 CN;

```

```
Initial Score      =      9  Optimized Score =      9  Significance =  4.69
Residue Identity  =    39%  Matches           =      9  Mismatches    =    14
Gaps              =      0  Conservative Substitutions =      0
```

X 10 20
TEFLSNYLTNVDDITLVPETLG
|| || || || | ||
100YNGILAS9GLKAFIGAMVNGMEYLQTFGEDVPYRRFPTLPAANFPNTERLYNKLTQDKELVVPSTP
820 830 840 850 860 870 880 890

X
R

VVKVGG
X

8. US-08-249-182-11 (1-23)

APCE FREDI PHYCOBILISOME 120 KD LINKER POLYPEPTIDE, CORE (L-C

```
ID  APCE_FREDI          STANDARD;          PRT; 1079 AA.
AC  P16566;
DT  01-AUG-1990 (REL. 15, CREATED)
DT  01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT  01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE  PHYCOBILISOME 120 KD LINKER POLYPEPTIDE, CORE (L-CM 92) (CORE-MEMBRANE
DE  LINKER PROTEIN).
GN  APCE.
OS  FREMYELLA DIPLOSIPHON (CALOTHRIX PCC 7601).
```

CC PROKARYOTA; GRACILICUTES; CATHYROTUBACTERIA;
 CC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90192765
 RA HOUMARD J., CAPUANO V., COLOMBANO M.V., COURSIN T., DE MARSAC N.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:2152-2156(1990).
 CC -!- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL
 CC ENERGY ACCEPTOR (BY ITS PHYCOBILIN-LIKE DOMAINS) AND AS A LINKER
 CC POLYPEPTIDE (BY ITS REPEATS AND ARMS) THAT STABILIZES THE
 CC PHYCOBILISOME CORE ARCHITECTURE.
 CC -!- SUBUNIT: PHYCOBILISOMES OF THIS ORGANISM ARE COMPOSED OF A TWO
 CC CYLINDER CORE, FROM WHICH SIX RODS RADIATE. THE CORE IS MAINLY
 CC COMPOSED OF ALLOPHYCOCYANIN ALPHA AND BETA CHAINS, AND OF THREE
 CC MINOR COMPONENTS: THE ALLOPHYCOCYANIN ALPHA-B CHAIN, A 18.3 KD
 CC POLYPEPTIDE, AND THE ANCHOR POLYPEPTIDE LCM.
 CC -!- SUBCELLULAR LOCATION: ANCHORS THE PHYCOBILISOME PERPENDICULARLY
 CC TO THE STROMAL SURFACE OF THE THYLAKOID MEMBRANE.
 CC -!- SIMILARITY: THE REPEATED DOMAINS ARE SIMILAR TO THE N-TERMINAL
 CC REGIONS OF PHYCOCYANIN ROD LINKER POLYPEPTIDES.
 CC -!- SIMILARITY: THE PHYCOBILIN-LIKE DOMAINS ARE SIMILAR TO PHYCOBILINS
 CC FROM VARIOUS SPECIES.
 DR EMBL; M20806; FDAPCA.
 DR PIR; A35088; A35088.
 KW PHYCOBILISOME; ELECTRON TRANSPORT; PHOTOSYNTHESIS; REPEAT.
 FT INIT_MET 0 0
 FT DOMAIN 17 75 PHYCOBILIN-LIKE.
 FT DOMAIN 76 143 PHYCOBILIN-LIKE LOOP.
 FT DOMAIN 144 236 PHYCOBILIN-LIKE.
 FT DOMAIN 237 284 ARM 1 (SPACING SEQUENCE).
 FT REPEAT 285 409 I.
 FT DOMAIN 410 546 ARM 2 (SPACING SEQUENCE).
 FT REPEAT 547 669 II.
 FT DOMAIN 670 743 ARM 3 (SPACING SEQUENCE).
 FT REPEAT 744 869 III.
 FT DOMAIN 870 953 ARM 4 (SPACING SEQUENCE).
 FT REPEAT 954 1079 IV.
 SQ SEQUENCE 1079 AA; 120325 MW; 5971605 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTVDDITLVPETLG
 || || || || || ||
 IRKYNQILATQGIRAFIGALVSSAEYAEVFGEDTPYRRYPTLPAANFPNTEKLYNQLTQNDLVVPSFKT
 830 840 850 860 870 X 880 890

X
 R
 VQPRLTLAGTSSSGRNGFTDLGRSSTSAQGQLGETANRCKPARIYRLSGTN
 900 910 920 930 940

9. US-08-249-182-11 (1-23)

BXD_CLOB0 BOTULINUM NEUROTOXIN TYPE D PRECURSOR (EC 3.4.24.-)

ID BXD_CLOB0 STANDARD; PRT; 1276 AA.
 AC P19321;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE BOTULINUM NEUROTOXIN TYPE D PRECURSOR (EC 3.4.24.-) (BONT/D).
 GN BOTD.

CC CLOSTRIDIUM BOTULINUM.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BVD/-3;
 RM 91016853
 RA BINZ T., KURAZONO H., POPOFF M.R., EKLUND M.W., SAKAGUCHI G.,
 RA KOZAKI S., KRIEGLSTEIN K., HENSCHEN A., GILL D.M., NIEMANN H.;
 RL NUCLEIC ACIDS RES. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CB16;
 RM 93042276
 RA SUNAGAWA H., OHYAMA T., WATANABE T., INOUE K.;
 RL J. VET. MED. SCI. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN=D-SA, AND D-1873;
 RM 89339741
 RA MORIISHI K., SYUTO B., KUBO S., OGUMA K.;
 RL INFECT. IMMUN. 57:2886-2891(1989).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOSOMAL-ASSOCIATED
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC A HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL
 CC ACTIVITY, WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE
 CC CHANNEL FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF BOTULINUM
 CC NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D STRAIN OF
 CC CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE BACTERIOPHAGE.
 CC -!- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
 CC NEUROTOXIN.
 CC -!- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
 CC REGION.
 DR EMBL; X54254; CBNTTD.
 DR EMBL; S49407; S49407.
 DR PIR; S11455; S11455.
 DR PROSITE; PS00142; ZINC_PROTEASE.
 KW NEUROTOXIN; TRANSMEMBRANE; HYDROLASE; METALLOPROTEASE; ZINC.
 FT CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
 FT CHAIN 443 1260 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
 FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 437 450 INTERCHAIN (PROBABLE).
 FT VARIANT 15 16 ND -> PV (IN STRAIN D-SA).
 FT VARIANT 17 18 ND -> LQ (IN STRAIN D-1873).
 FT VARIANT 452 452 K -> Q (IN STRAIN D-SA).
 FT VARIANT 457 457 R -> T (IN STRAIN D-SA).
 FT VARIANT 457 457 R -> F (IN STRAIN D-1873).
 FT VARIANT 462 462 A -> D (IN STRAIN D-1873).
 FT VARIANT 489 489 K -> N (IN STRAIN CB16).
 FT VARIANT 644 644 N -> K (IN STRAIN CB16).
 FT VARIANT 1122 1122 Q -> R (IN STRAIN CB16).
 SQ SEQUENCE 1276 AA; 146871 MW; 8705534 CN;

Initial Score = 9 Optimized Score = 9 Significance = 4.69
 Residue Identity = 39% Matches = 9 Mismatches = 14
 Gaps = 0 Conservative Substitutions = 0

X 10 EV
 TEFLSNVLTNVDDITLVPETLG
 ||| | |||
 DGQVPINPEIVDPLLPNVNHEPLNLPGEIEVFYDDITKYVDYLNSEYYLESSQKLSNNVENITLTTTSVEEALG
 500 510 520 530 540 550 560

X
R

YSNKIYTFPLSLAEKVNKGVOAGLFLNWANEVVEDFTTNIMKKDTLDKISD
 570 580 590 600 610 620

10. US-08-249-182-11 (1-23)

VNS4_CVPRM NONSTRUCTURAL PROTEIN 4 (X1 PROTEIN) (ORF 3).

ID VNS4_CVPRM STANDARD; PRT; 82 AA.
 AC P24415;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE NONSTRUCTURAL PROTEIN 4 (X1 PROTEIN) (ORF 3).
 GN NS4.
 OS PORCINE RESPIRATORY CORONAVIRUS (STRAIN RM4) (PRCV), AND PORCINE
 OS RESPIRATORY CORONAVIRUS (STRAIN 86/137004 / BRITISH ISOLATE) (PRCV).
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; CORONAVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RM4;
 RM 91073120
 RA RASSCHAERT D., DUARTE M., LAUDE H.;
 RL J. GEN. VIROL. 71:2599-2607(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=86/137004;
 RM 92116634
 RA BRITTON P., MAWDITT K.L., PAGE K.W.;
 RL VIRUS RES. 21:181-198(1991).
 DR EMBL; X60056; PRCVORFS.
 DR PIR; C36607; C36607.
 DR PIR; S21309; S21309.
 DR PIR; S24280; S24280.
 KW NONSTRUCTURAL PROTEIN.
 SQ SEQUENCE 82 AA; 9268 MW; 34513 CN;

Initial Score = 8 Optimized Score = 8 Significance = 3.91
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

X 10 20 X
 TEFLSNVLTNVDDITLVPETLGR
 || || | |||
 MTFPRALTVIDDNGMVISIIFWFLIIILILLSIALLNIIKLCMVCCNLGRTVIIIPVQHAYDAYKNFMRIK
 10 20 30 40 50 60 70
 AYNPDGALLV
 80

11. US-08-249-182-11 (1-23)

VNS4_CVPPU NONSTRUCTURAL PROTEIN 4 (X1 PROTEIN) (ORF 3).

ID VNS4_CVPPU STANDARD; PRT; 82 AA.
 AC P09048;
 DT 01-NOV-1988 (REL. 09, CREATED)
 DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)

DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE NONSTRUCTURAL PROTEIN 4 (X1 PROTEIN) (ORF 3).
 GN NS4.
 OS PORCINE TRANSMISSIBLE GASTROENTERITIS CORONAVIRUS (STRAIN PURDUE).
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; CORONAVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 88078100
 RA RASSCHAERT D., GELFI J., LAUDE H.;
 RL BIOCHIMIE 69:591-600(1987).
 DR EMBL; X06371; COTGEV3.
 DR PIR; S01741; S01741.
 DR PIR; S04891; S04891.
 KW NONSTRUCTURAL PROTEIN.
 SQ SEQUENCE 82 AA; 9239 MW; 31619 CN;

Initial Score = 8 Optimized Score = 8 Significance = 3.91
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

			X		10		20	X	
			TEFLSNYLTNVDDITLVPETLGR						
MTFPRALTVIDDNGMVINIIFWFLIIILILLSIALLNIIKLCMVCCNLGRTVIIVPAQHAYDAYKNFMRIK									
	10	20	30	40	50	60	70		
AYNPDGALLA									
80									

12. US-08-249-182-11 (1-23)

REPB_STRPN REPLICATION PROTEIN REPB.

ID REPB_STRPN STANDARD; PRT; 210 AA.
 AC P13921;
 DT 01-JAN-1990 (REL. 13, CREATED)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
 DE REPLICATION PROTEIN REPB.
 GN REPB.
 OS STREPTOCOCCUS PNEUMONIAE, AND STREPTOCOCCUS AGALACTIAE.
 OG PLASMID PLS1, AND PLASMID PMV158.
 OC PROKARYOTA; FIRMICUTES; COCCI; STREPTOCOCCAEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PLASMID=PLS1;
 RM 87226167
 RA LACKS S.A., LOPEZ P., GREENBERG B., ESPINOSA M.;
 RL J. MOL. BIOL. 192:753-765(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC PLASMID=PMV158;
 RM 90016790
 RA VAN DER LELIE D., BRON S., VENEMA G., OSKAM L.;
 RL NUCLEIC ACIDS RES. 17:7283-7294(1989).
 CC -!- FUNCTION: IS ESSENTIAL FOR PLASMID REPLICATION. NICKS THE POSITIVE
 CC STRAND AT THE PLUS ORIGIN OF REPLICATION.
 CC -!- SIMILARITY: WITH REPLICATION PROTEINS FROM OTHER GRAM+ BACTERIAL
 CC PLASMIDS. REPLICATING WITH THE ROLLING-CIRCLE MECHANISM.
 DR EMBL; X15669; SAREPAB.
 DR EMBL; M29725; PPCG1.
 DR PIR; B25599; B25599.
 DR PIR; S05981; S05981.
 KW PLASMID; DNA REPLICATION; TOPOISOMERASE.
 SQ SEQUENCE 210 AA; 24250 MW; 229833 CN;

Initial Score = 8 Optimized Score = 8 Significance = 3.91
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

      X      10      20  X
      TEFLSNYLTVDDITLVPETLGR
          ||      |  |  |||
MAKEKARYFTFLLYPESIPSDWELKLETLGVPMAISPLHDKDKSSIKGQKYKKAHVHLYIAKNPVTADSVR
      10      20      30      40      50      60      70

KKIKLLLGE
      80
```

13. US-08-249-182-11 (1-23)
HYPE_BRAJA HYPE PROTEIN.

ID HYPE_BRAJA STANDARD; PRT; 321 AA.
AC P31906;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE HYPE PROTEIN.
GN HYPE.
OS BRADYRHIZOBIUM JAPONICUM.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC RHIZOBIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB1809;
RM 93287991
RA VAN SOOM C., VERRETH C., SAMPAIO M.J., VANDERLEYDEN J.;
RL MOL. GEN. GENET. 239:235-238(1993).
CC -!- SIMILARITY: BELONGS TO THE HYPE FAMILY.
DR EMBL; Z17373; BJHYPHOX.
DR PIR; S28642; S28642.
SQ SEQUENCE 321 AA; 33581 MW; 463510 CN;

Initial Score = 8 Optimized Score = 8 Significance = 3.91
Residue Identity = 34% Matches = 8 Mismatches = 15
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      20
                                TEFLSNYLTVDDITLVPETLG
                                    ||      |  |  |||
FADLKLIAESMGAAAREADVHIITGDTKVVVERGKADGLFISTAGVGVVPDGLDLSAEKARVGDRLISGTLG
      90      100      110      120      130      140      150

X
R

DHGVAIMSKRQNLAFETEIVSDSASLHDLVARMVQAGGRGIRLMRDPTRGG
160      170      180      190      200
```

14. US-08-249-182-11 (1-23)
EGF2_STRPU EPIDERMAL GROWTH FACTOR-RELATED PROTEIN PRECURSOR

ID EGF2_STRPU STANDARD; PRT; 325 AA.
AC P15216;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE EPIDERMAL GROWTH FACTOR-RELATED PROTEIN PRECURSOR (SPEGF2).
OS STRONGYLOCENTROTUS PURPURATUS (PURPLE SEA URCHIN).
OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA;

```

UC      EVECCHINDIDEA.
RN      [1]
RP      SEQUENCE FROM N.A.
RM      90049203
RA      YANG D., ANGERER L.M., ANGERER R.C.;
RL      SCIENCE 246:806-808(1989).
CC      -!- SIMILARITY: THE PROTEIN INCLUDES 4 EGF-LIKE REPEATS.
CC      -!- SIMILARITY: EACH OF THE EGF-LIKE REPEAT IS VERY SIMILAR TO THE
CC          EXOGASTRULA-INDUCING PEPTIDES FROM THE SEA URCHIN ANTHOCIDARIS
CC          CRASSISPINA.
DR      EMBL; M29004; SPEGF2.
KW      EGF-LIKE DOMAIN; REPEAT; SIGNAL.
FT      SIGNAL          1          17          POTENTIAL.
FT      CHAIN           18         325          EPIDERMAL GROWTH FACTOR-RELATED PROTEIN.
FT      DOMAIN          47         311          4 X EGF-TYPE REPEATS.
FT      REPEAT          47         105          EGF-LIKE 1.
FT      REPEAT          106        176          EGF-LIKE 2.
FT      REPEAT          177        248          EGF-LIKE 3.
FT      REPEAT          249        311          EGF-LIKE 4.
SQ      SEQUENCE      325 AA;  36898 MW;  486695 CN;

```

```

                                     X      10      20
                                     TEFLSNYLTNVDDITLVPETLG
                                     ||||      |      ||
NRCLSDTSNCDGHGICQLSTFGRNERYICFCALGFRNNYGGCSPYTPREIEFLSYVARDLEEMLTRDSLGL
110      120      130      140      150      160      170      180

X
R
|
RCKSDTHNCDEAGQCVTKTYGRYAGEYICVCNHGYRNNAYGGCSPMTTREI
X      190      200      210      220      230

```

```

ID      MYB_AVIMB          STANDARD;          PRT;    382 AA.
AC      P01104;
DT      21-JUL-1986 (REL. 01, CREATED)
DT      01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT      01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE      MYB TRANSFORMING PROTEIN.
CN      V-MYB.
OS      AVIAN MYELOBLASTOSIS VIRUS.
OC      VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC      ONCOVIRINAE.
RN      [1]
RP      SEQUENCE FROM N.A.
RM      83129359
RA      KLEMPNAUER K.-H., GONDA T.J., BISHOP J.M.;
RL      CELL 31:453-463(1982).
RN      [2]
RP      SEQUENCE FROM N.A.
RM      82223743
RA      RUSHLOW K.E., LAUTENBERGER J.A., PAPAS T.S., BALUDA M.A., PERBAL B.,
RA      CHIRIKJIAN J.G., REDDY E.P.;
RL      SCIENCE 216:1421-1423(1982).
CC      -!- FUNCTION: MYB IS A DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZE
CC      THE SEQUENCE YAAC(G/T)G.
CC      -!- DISEASE: THE V-MYB ONCOGENE TRANSFORMS IMMATURE MYELOMONOCYTIC
CC      AVIAN CELLS IN CULTURE AND INDUCES MYELOBLASTOSIS (MYELOID

```


CC LEUKEMIA/ IN CHICKENS.
 DR EMBL; J02012; REONVMYB.
 DR PIR; A01347; QDVV.
 DR TFD; P00028; RELEASE 3.0.
 DR PROSITE; PS00037; MYB_1.
 DR PROSITE; PS00334; MYB_2.
 KW TRANSFORMING PROTEIN; ONCOGENE; NUCLEAR PROTEIN; DNA-BINDING; REPEAT.
 FT DNA_BIND 1 15 MYB (PARTIAL).
 FT DNA_BIND 16 67 MYB.
 FT DNA_BIND 68 118 MYB.
 SQ SEQUENCE 382 AA; 43061 MW; 744659 CN;

Initial Score = 8 Optimized Score = 8 Significance = 3.91
 Residue Identity = 34% Matches = 8 Mismatches = 15
 Gaps = 0 Conservative Substitutions = 0

X 10 20
 TEFLSNYLTVDDITLVPETLG
 || || || ||
 WHSTTVADNTRTSGDNAPVSCLEHHCTSPSPVDHGCLEESASPARCMIVHGSNILDNVKNLLEFAETLQ
 260 270 280 290 300 X 310 320

X
 R

LIDSFLNTSSNHNENLNDNPALTSTPVCCHKMSVTTPFHKDQTFTEYRKMH
 330 340 350 360 370

> 0 <
 0| 0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences

Release 5.4 *Seq. 1*

Results file u249_1a.res made by on Thu 22 Sep 94 10:06:41-PDT.

Query sequence being compared:US-08-249-182-1 (1-5)

Number of sequences searched: 42145

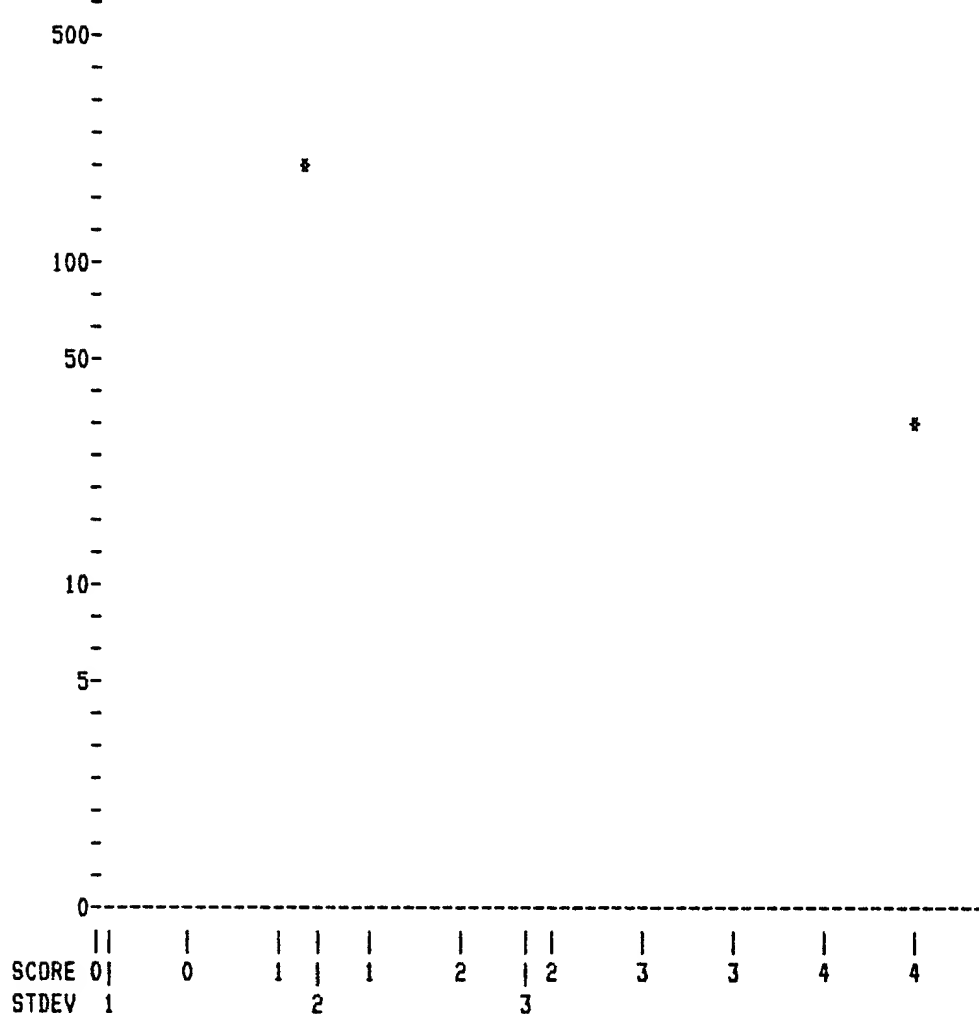
Number of scores above cutoff: 4616

Results of the initial comparison of US-08-249-182-1 (1-5) with:
 Data bank : A-GeneSeq 15, all entries

100000-
 -
 N -
 U50000-
 M -
 B *
 E -
 R -
 -
 O -
 F10000-
 -
 S -
 E 5000-
 Q -
 U -
 E -
 N -
 C -
 E -
 S 1000-
 -

*

*



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.02

Times:	CPU	Total Elapsed
	00:00:26.99	00:00:37.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	4616

Cut-off raised to 2.
Cut-off raised to 3.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Init. Opt.		Sig.	Frame
		Length	Score		
**** 3 standard deviations above mean ****					
1. R37443	Autotaxin peptide ATX 18.	6	4	4	3.92 0
2. P71436	Immunomodulator peptide #12 i	9	4	4	3.92 0
3. P71435	Immunomodulator peptide #11 i	9	4	4	3.92 0
4. P71423	Immunomodulator peptide #2 in	21	4	4	3.92 0
5. P71422	Immunomodulator peptide #1 in	21	4	4	3.92 0
6. P71493	Antigenic peptide cross-react	32	4	4	3.92 0
7. R42697	p16 of nef peptide of HIV-1.	35	4	4	3.92 0
8. P90958	Peptide corresp. to AA171-205	35	4	4	3.92 0
9. P92270	Peptide PF16 from HIV-1 prote	37	4	4	3.92 0
10. R42698	p17 of nef peptide of HIV-1.	65	4	4	3.92 0
11. P92271	Peptide PF17 from HIV-1 prote	67	4	4	3.92 0
12. R11188	Wheat 33kD protein pre-sequen	105	4	4	3.92 0
13. R38893	Nef protein of HIV-1.	206	4	4	3.92 0
14. P61515	Sequence of E' protein.	206	4	4	3.92 0
15. P60423	Sequence of LAV virus ORF F p	216	4	4	3.92 0
16. P70863	Sequence of S2 subunit of Bor	226	4	4	3.92 0
17. P96014	Pertussis toxin subunit S2.	226	4	4	3.92 0
18. R23966	Protein sequence of plasmid p	234	4	4	3.92 0
19. R45021	Staphylococcal enterotoxin ET	246	4	4	3.92 0
20. P92062	Sequence of Isopenicillin N s	333	4	4	3.92 0
21. R20006	Zonula occludens toxin.	399	4	4	3.92 0
22. R44220	Threonine synthase.	481	4	4	3.92 0
23. P81325	Threonine synthetase.	495	4	4	3.92 0
24. R10274	Simian immunodeficiency viru	502	4	4	3.92 0
25. R32356	Excitory amino acid receptor	956	4	4	3.92 0
26. R27931	SPS protein.	1068	4	4	3.92 0
27. R20198	Sucrose phosphate synthase fr	1068	4	4	3.92 0
28. R24033	Soluble mannose receptor pept	1456	4	4	3.92 0
29. R44432	eryA region polypeptide modul	2986	4	4	3.92 0
30. R10834	Rianodin receptor.	4987	4	4	3.92 0
31. R11510	Ryanodine receptor deduced fr	5072	4	4	3.92 0
**** 2 standard deviations above mean ****					
32. R28702	BAG 75 50 kD fragment N-termi	6	3	3	2.94 0
33. R14168	ACE-inhibiting hexapeptide, I	6	3	3	2.94 0
34. R43915	Pyruvate kinase conserved dom	8	3	3	2.94 0
35. R07942	TNF binding protein fraction	9	3	3	2.94 0
36. R44069	Pulmonary surfactant protein	10	3	3	2.94 0
37. R44065	Pulmonary surfactant protein	13	3	3	2.94 0
38. R28701	BAG 75 N-terminal.	13	3	3	2.94 0
39. P70837	Sequence encoded by the Beta-	14	3	3	2.94 0
40. R34888	Human TSH residues 61-75.	15	3	3	2.94 0

1. US-08-249-182-1 (1-5)

R37443 Autotaxin peptide ATX 18.

ID R37443 standard; peptide; 6 AA.
AC R37443;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 18.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

P1 AF012534 N; Liotta LA; Schiffman E; Stracke R.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 18. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 6 AA;
 SQ 2 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 1 H;
 SQ 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 0 S; 0 T; 1 W; 0 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 WHVAAN
 X X

2. US-08-249-182-1 (1-5)

P71436 Immunomodulator peptide #12 inhibits HIV-T4 intera

ID P71436 standard; Protein; 9 AA.
 AC P71436;
 DT 03-MAY-1991 (first entry)
 DE Immunomodulator peptide #12 inhibits HIV-T4 interaction.
 KW AIDS; T4 cell receptor; immunomodulation.
 OS Synthetic.
 PN W08703601-A.
 PD 18-JUN-1987.
 PF 08-DEC-1986; 402717.
 PR 06-DEC-1985; FR-018155.
 PA (INSP) INST PASTEUR.
 PA (AUFR/) AUFRAY C.
 PI Auffray C, Montagnier L, Klatzmann D, Charron D;
 DR WPI; 87-177935/25.
 PT New peptide derivs. contg. specified exposed tetra:peptide
 PT sequences - inhibiting interaction of AIDS virus with T4 cell
 PT receptors
 PS Claim 18; Page 50; 57pp; French.
 CC The peptide corresponds to the conserved sequence immediately
 CC after the RFDS peptide motif of the Type II HLA antigen of ARV2.
 CC It is used to produce monoclonal antibodies specific to the peptide.
 CC See also P71422-P71435 and P71437.
 SQ Sequence 9 AA;
 SQ 2 A; 1 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 1 L; 1 K; 0 M; 1 F; 0 P; 0 S; 0 T; 0 W; 0 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 KLAFHHVAR
 X X

3. US-08-249-182-1 (1-5)

P71435 Immunomodulator peptide #11 inhibits HIV-T4 intera

ID P71435 standard; Protein; 9 AA.
AC P71435;
DT 03-MAY-1991 (first entry)
DE Immunomodulator peptide #11 inhibits HIV-T4 interaction.
KW AIDS; T4 cell receptor; immunomodulation.
OS Synthetic.
PN W08703601-A.
PD 18-JUN-1987.
PF 08-DEC-1986; 402717.
PR 06-DEC-1985; FR-018155.
PA (INSP) INST PASTEUR.
PA (AUFR/) AUFRA Y C.
PI Auffray C, Montagnier L, Klatzmann D, Charron D;
DR WPI; 87-177935/25.
PT New peptide derivs. contg. specified exposed tetra:peptide
PT sequences - inhibiting interaction of AIDS virus with T4 cell
PT receptors
PS Claim 18; Page 50; 57pp; French.
CC The peptide corresponds to the conserved sequence immediately
CC after the RFDS peptide motif of the Type II HLA antigens of
CC LV and LAV1a. It is used to produce monoclonal antibodies specific
CC to the peptide.
CC See also P71422-P71434 and P71436-7.
SQ Sequence 9 AA;
SQ 2 A; 2 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 2 H;
SQ 0 I; 1 L; 0 K; 0 M; 1 F; 0 P; 0 S; 0 T; 0 W; 0 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
RLAFHHVAR
X X

4. US-08-249-182-1 (1-5)

P71423 Immunomodulator peptide #2 inhibits HIV-T4 interac

ID P71423 standard; Protein; 21 AA.
AC P71423;
DT 03-MAY-1991 (first entry)
DE Immunomodulator peptide #2 inhibits HIV-T4 interaction.
KW AIDS; T4 cell receptor; immunomodulation.
OS Synthetic.
PN W08703601-A.
PD 18-JUN-1987.
PF 08-DEC-1986; 402717.
PR 06-DEC-1985; FR-018155.
PA (INSP) INST PASTEUR.
PA (AUFR/) AUFRA Y C.
PI Auffray C, Montagnier L, Klatzmann D, Charron D;
DR WPI; 87-177935/25.
PT New peptide derivs. contg. specified exposed tetra:peptide
PT sequences - inhibiting interaction of AIDS virus with T4 cell
PT receptors
PS Claim 5; Page 48; 57pp; French.
CC The peptide is a specific example of a peptide comprising the
CC tetrapeptide motif RFDS (pref. at position 7 to 10 and optionally

CC having RE at positions 1 and 2 and/or EL at positions 20 and 21).
 CC It interferes with interaction between the AIDS virus and T4
 CC receptors on lymphocytes. The peptide also has immunomodulatory
 CC activity. It is useful in diagnosis to detect antibodies to the
 CC region of the viral genome containing the RFDS sequence.
 CC See also P71422 and P71424-P71437.
 SQ Sequence 21 AA;
 SQ 2 A; 3 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 3 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 3 L; 1 K; 0 M; 2 F; 0 P; 1 S; 0 T; 1 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 HHVAR
 ||||
 REVLEWRFD SKLAFHHVAREL
 10 X 20

5. US-08-249-182-1 (1-5)
 P71422 Immunomodulator peptide #1 inhibits HIV-T4 interac

ID P71422 standard; peptide; 21 AA.
 AC P71422;
 DT 03-MAY-1991 (first entry)
 DE Immunomodulator peptide #1 inhibits HIV-T4 interaction.
 KW AIDS; T4 cell receptor; immunomodulation.
 OS Synthetic.
 PN W08703601-A.
 PD 18-JUN-1987.
 PF 08-DEC-1986; 402717.
 PR 06-DEC-1985; FR-018155.
 PA (INSP) INST PASTEUR.
 PA (AUFR/) AUFRA Y C.
 PI Auffray C, Montagnier L, Klatzmann D, Charron D;
 DR WPI; 87-177935/25.
 PT New peptide derivs. contg. specified exposed tetra:peptide
 PT sequences - inhibiting interaction of AIDS virus with T4 cell
 PT receptors
 PS Claim 5; Page 48; 57pp; French.
 CC The peptide is a specific example of a peptide comprising the
 CC tetrapeptide motif RFDS (pref. at position 7 to 10 and optionally
 CC having RE at positions 1 and 2 and/or EL at positions 20 and 21).
 CC It interferes with interaction between the AIDS virus and T4
 CC receptors on lymphocytes. The peptide also has immunomodulatory
 CC activity. It is useful in diagnosis to detect antibodies to the
 CC region of the viral genome containing the RFDS sequence.
 CC See also P71423-P71437.
 SQ Sequence 21 AA;
 SQ 2 A; 4 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 3 E; 0 Z; 0 G; 2 H;
 SQ 0 I; 3 L; 0 K; 0 M; 2 F; 0 P; 1 S; 0 T; 1 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 HHVAR
 ||||
 REVLEWRFD SRLAFHHVAREL
 10 X 20

6. US-08-249-182-1 (1-5)

ID P71493 standard; peptide; 32 AA.
AC P71493;
DT 01-MAY-1991 (first entry)
DE Antigenic peptide cross-reactive with HTLV-III env protein 3'ORF.
KW Human T-cell Lymphotropic virus; p27 antigen.
OS Human t-cell lymphotropic virus.
PN W08702988-A.
PD 21-MAY-1987.
PF 05-NOV-1986; U02381.
PR 07-NOV-1985; US-795997.
PA (HARD) HARVARD COLLEGE.
PI Essex ME, Allan JS, Lee TH;
DR WPI; 87-150611/21.
PT Human T-cell lymphotropic virus protein - used for assaying for
PT antibodies and in assaying for antigenic determinants for
PT predicting the course of HTLV-III infection
PS Claim 7; Page 14; 25pp; English.
CC The p27 antigen has an determinant cross-reactive with the protein
CC encoded by the ORF 3'to the env gene of HTLV-III. It may be used in
CC diagnosis and monitoring of the course of the disease.
SQ Sequence 32 AA;
SQ 2 A; 4 R; 1 N; 3 D; 0 B; 1 C; 0 Q; 5 E; 0 Z; 0 G; 3 H;
SQ 0 I; 3 L; 1 K; 0 M; 2 F; 2 P; 1 S; 0 T; 1 W; 1 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
DDPEREVLEWRDSRLAFHHVARELHPEYFKNC
10 20 X 30

7. US-08-249-182-1 (1-5)

R42697 p16 of nef peptide of HIV-1.

ID R42697 standard; protein; 35 AA.
AC R42697;
DT 10-NOV-1993 (first entry)
DE p16 of nef peptide of HIV-1.
KW AIDS; antibody; p25; gp110; gp41; assay; detection;
KW immunity; vaccine.
OS Human immunodeficiency virus-1.
FH Key Location/Qualifiers
FT Modified_site 35
FT /note= "Cys(acetanidomethyl)"
PN US5221610-A.
PD 22-JUN-1993.
PF 26-MAY-1988; 199143.
PR 26-MAY-1988; US-199143.
PR 04-SEP-1991; US-754300.
PA (INRM) INST NAT SANTE & RECH MEDICALE.
PA (INSP) INST PASTEUR.
PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;
PI Rietschoten JV, Rochat H, Sabatier JM;
DR WPI; 93-213434/26.
PT Diagnosis of HIV infection - by detecting HIV antibodies using
PT antigenic polypeptide derived from nef protein of HIV-1
PS Disclosure; Page 3; 15pp; English.
CC The peptide is expressed in vivo in HIV infected patients before
CC detectable ants. of p25, gp110 and gp41 are expressed. Thus, it
CC can be used in assays for early detection of HIV.

CC It can also be used to raise antibodies for use in detection,
 CC to induce cellular immunity or to raise neutralising antibodies
 CC that either inactivate the AIDS virus or reduce the viability of
 CC the virus in vivo or destroy infected cells.
 CC The peptide may be used in viral vaccines.
 SQ Sequence 35 AA;
 SQ 2 A; 4 R; 1 N; 3 D; 0 B; 1 C; 0 Q; 5 E; 0 Z; 1 G; 3 H;
 SQ 0 I; 3 L; 1 K; 1 M; 3 F; 2 P; 1 S; 0 T; 1 W; 1 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 GMDDPEREVLEWRFSRLAFHHVARELHPEYFKNC
 10 20 X 30

8. US-08-249-182-1 (1-5)
 P90958 Peptide corresp. to AA171-205 of HIV-1 protein F f

ID P90958 standard; peptide; 35 AA.
 AC P90958;
 DT 23-FEB-1990 (first entry)
 DE Peptide corresp. to AA171-205 of HIV-1 protein F fragment
 KW HIV
 OS
 PN W08909227-A.
 PD 05-OCT-1989.
 PF 31-MAR-1989; F00151.
 PR 01-APR-1988; FR-004405.
 PA (INSP) Inst Pasteur.
 PI Montagnier L, Rochat H, Bahraoui EM, Chamaret S;
 DR WPI; 89-309503/42.
 PT New human immuno-deficiency virus peptide(s) - corresp. to protein F
 PT fragments, useful for acquired immunodeficiency syndrome diagnosis
 PS Claim 3; page 43; 57pp; French.
 CC This can be used as an immunoassay reagent for AIDS diagnosis, in
 CC treating AIDS and in the prodn. of other peptides and vaccines. They are
 CC prep'd. by conventional synthesis.
 SQ Sequence 35 AA;
 SQ 2 A; 4 R; 1 N; 3 D; 0 B; 1 C; 0 Q; 5 E; 0 Z; 1 G; 3 H;
 SQ 0 I; 3 L; 1 K; 1 M; 3 F; 2 P; 1 S; 0 T; 1 W; 1 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 GMDDPEREVLEWRFSRLAFHHVARELHPEYFKNC
 10 20 X 30

9. US-08-249-182-1 (1-5)
 P92270 Peptide PF16 from HIV-1 protein F fragment

ID P92270 standard; peptide; 37 AA.
 AC P92270;
 DT 23-FEB-1990 (first entry)
 DE Peptide PF16 from HIV-1 protein F fragment
 KW HIV
 OS

FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note="OH, NH2 or 1-5 AA"
 FT Misc-difference 37
 FT /note="OH, NH2 or 1-5 AA"
 PN W08909227-A.
 PD 05-OCT-1989.
 PF 31-MAR-1989; F00151.
 PR 01-APR-1988; FR-004405.
 PA (INSP) Inst Pasteur.
 PI Montagnier L, Rochat H, Bahraoui EM, Chamaret S;
 DR WPI; 89-309503/42.
 PT New human immuno-deficiency virus peptide(s) - corresp. to protein F
 PT fragments, useful for acquired immunodeficiency syndrome diagnosis
 PS Claim 4; page 43; 57pp; French.
 CC This can be used as an immunoassay reagent for AIDS diagnosis, in
 CC treating AIDS and in the prodn. of other peptides and vaccines. It is
 CC prepd. by conventional synthesis.
 SQ Sequence 37 AA;
 SQ 2 A; 4 R; 1 N; 3 D; 0 B; 1 C; 0 Q; 5 E; 1 Z; 1 G; 3 H;
 SQ 0 I; 3 L; 1 K; 1 M; 3 F; 2 P; 1 S; 0 T; 1 W; 1 Y; 2 V;
 SQ 1 Others;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 XGMDDPEREVLEWRFD SRLAFHHVARELHPEYFKNCZ
 10 20 X X 30

10. US-08-249-182-1 (1-5)

R42698 p17 of nef peptide of HIV-1.

ID R42698 standard; protein; 65 AA.
 AC R42698;
 DT 10-NOV-1993 (first entry)
 DE p17 of nef peptide of HIV-1.
 KW AIDS; antibody; p25; gp110; gp41; assay; detection;
 KW immunity; vaccine.
 OS Human immunodeficiency virus-1.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "Cys(acetamidomethyl)"
 FT Modified_site 65
 FT /note= "Cys(acetamidomethyl)"
 PN US5221610-A.
 PD 22-JUN-1993.
 PF 26-MAY-1988; 199143.
 PR 26-MAY-1988; US-199143.
 PR 04-SEP-1991; US-754300.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;
 PI Rietschoten JV, Rochat H, Sabatier JM;
 DR WPI; 93-213434/26.
 PT Diagnosis of HIV infection - by detecting HIV antibodies using
 PT antigenic polypeptide derived from nef protein of HIV-1
 PS Disclosure; Page 3; 15pp; English.
 CC The peptide is expressed in vivo in HIV infected patients before
 CC detectable ants. of p25, gp110 and gp41 are expressed. Thus, it
 CC can be used in assays for early detection of HIV.
 CC It can also be used to raise antibodies for use in detection,

CC to induce cellular immunity or to raise neutralizing antibodies
 CC that either inactivate the AIDS virus or reduce the viability of
 CC the virus in vivo or destroy infected cells.
 CC The peptide may be used in viral vaccines.
 SQ Sequence 65 AA;
 SQ 3 A; 4 R; 3 N; 4 D; 0 B; 2 C; 0 Q; 9 E; 0 Z; 2 G; 5 H;
 SQ 0 I; 7 L; 4 K; 1 M; 3 F; 5 P; 3 S; 1 T; 1 W; 2 Y; 6 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 CYKLVPVEPDKVEEANKGENTSLHPVSLHGMDPPEREVLEWRFD SRLAFHHVARELHPEYFKNC
 10 20 30 40 50 X 60

11. US-08-249-182-1 (1-5)

P92271 Peptide PF17 from HIV-1 protein F fragment

ID P92271 standard; peptide; 67 AA.
 AC P92271;
 DT 23-FEB-1990 (first entry)
 DE Peptide PF17 from HIV-1 protein F fragment
 KW HIV
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note="OH, NH2 or 1-5 AA"
 FT Misc-difference 67
 FT /note="OH, NH2 or 1-5 AA"
 PN W08909227-A.
 PD 05-OCT-1989.
 PF 31-MAR-1989; F00151.
 PR 01-APR-1988; FR-004405.
 PA (INSP) Inst Pasteur.
 PI Montagnier L, Rochat H, Bahraoui EM, Chamaret S;
 DR WPI; 89-309503/42.
 PT New human immuno-deficiency virus peptide(s) - corresp. to protein F
 PT fragments, useful for acquired immunodeficiency syndrome diagnosis
 PS Claim 4; page 43; 57pp; French.
 CC This can be used as an immunoassay reagent for AIDS diagnosis, in
 CC treating AIDS and in the prodn. of other peptides and vaccines. It is
 CC prepd. by conventional synthesis.
 SQ Sequence 67 AA;
 SQ 3 A; 4 R; 3 N; 4 D; 0 B; 2 C; 0 Q; 9 E; 1 Z; 2 G; 5 H;
 SQ 0 I; 7 L; 4 K; 1 M; 3 F; 5 P; 3 S; 1 T; 1 W; 2 Y; 6 V;
 SQ 1 Others;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 CYKLVPVEPDKVEEANKGENTSLHPVSLHGMDPPEREVLEWRFD SRLAFHHVARELHPEYFKNCZ
 10 20 30 40 50 X X 60

12. US-08-249-182-1 (1-5)

R11188 Wheat 33kD protein pre-sequence and N-terminal lin

ID R11188 standard; Protein; 105 AA.

DT 22-MAY-1991 (first entry)
 DE Wheat 33kD protein pre-sequence and N-terminal linker.
 KW wheat; 33kDa protein; photosynthetic oxygen-evolving complex;
 KW thylakoid lumen.
 OS Triticum aestivum.
 FH Key Location/Qualifiers
 FT Peptide 1..79
 FT /label= 33kDa protein pre-sequence
 FT /note= "targets proteins to thylakoid lumen"
 PN W09102800-A.
 PD 07-MAR-1991.
 PF 14-AUG-1990; G01281.
 PR 14-AUG-1989; GB-018496.
 PA (AGRI-) AGRIC GENETICS CO LTD.
 PA (ADTE-) ADVANCED TECHNOLOGIES (CAMBRIDGE) LTD.
 PA (BIOT-) BIOTAL LTD.
 PA (BRPE) BP NUTRITION LTD.
 PA (CIBA) CIBA-GEIGY PLC.
 PA (ICIL) IMPERIAL CHEMICAL INDUSTRIES PLC.
 PA (RHON) RHONE-POULENC LTD.
 PA (SCHD) SCHERING AGROCHEMICALS LTD.
 PA (SHEL) SHELL RESEARCH LTD.
 PA (UNIL) UNILEVER UK.
 PI Robinson C.
 DR WPI; 91-087283/12.
 DR Q-PSDB; 011022.
 PT Novel chimeric gene used for transforming plant cell - useful for
 PT delivering heterologous passenger protein into thylakoid lumen of
 PT chloroplast.
 PS Claim 5; Page 17; 19pp; English.
 CC The sequence comprises the 79 amino acid pre-sequence of the 33kDa
 CC protein of the photosynthetic oxygen-evolving complex of wheat,
 CC directly linked to the first 26 N-terminal amino acids of the mature
 CC wheat 33kDa protein. The latter constitutes an optional linker
 CC between the pre-sequence and an heterologus protein sequence.
 CC Vectors containing a chimaeric gene of the invention (i.e. the pre-
 CC sequence, optional linker and heterologous protein-encoding
 CC sequences) can be used to transform a plant cell for production of
 CC the heterologous protein which is targetted to the thylakoid lumen
 CC of the chloroplast.
 SQ Sequence 105 AA;
 SQ 24 A; 6 R; 0 N; 4 D; 0 B; 2 C; 3 Q; 4 E; 0 Z; 9 G; 1 H;
 SQ 4 I; 7 L; 6 K; 4 M; 3 F; 3 P; 11 S; 8 T; 0 W; 1 Y; 5 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

          X  X
        WHVAR
          ||||
MAASLQAAATLMPAKIGGRASSARPSHVARAFGV DAGARITCSLQSDIREVASKCADA AKMAGFALATSAL
      10      20      X 30      40      50      60      70

LVSGATAEG
      80
  
```

13. US-08-249-182-1 (1-5)
 R38893 Nef protein of HIV-1.

ID R38893 standard; Protein; 206 AA.
 AC R38893;
 DT 10-NOV-1993 (first entry)
 DE Nef protein of HIV-1.

KW AIDS; antibody; p25; gp110; gp41; assay; detection;
 OS Human immunodeficiency virus-1.
 PN US5221610-A.
 PD 22-JUN-1993.
 PF 26-MAY-1988; 199143.
 PR 26-MAY-1988; US-199143.
 PR 04-SEP-1991; US-754300.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 PI Bahraoui EM, Chamaret S, Ferris S, Granier C, Montagnier L;
 PI Rietschoten JV, Rochat H, Sabatier JM;
 DR WPI; 93-213434/26.
 PT Diagnosis of HIV infection - by detecting HIV antibodies using
 PT antigenic polypeptide derived from nef protein of HIV-1
 PS Disclosure; Fig 2; 15pp; English.
 CC The nef protein comprises peptides which are expressed in vivo in HIV
 CC infected patients before detectable ants. of p25, gp110 and gp41 are
 CC expressed. Thus, they can be used in assays for early detection of HIV.
 CC They can also be used to raise antibodies for use in detection,
 CC to induce cellular immunity or to raise neutralising antibodies
 CC that either inactivate the AIDS virus or reduce the viability of
 CC the virus in vivo or destroy infected cells.
 CC The peptides may be used in viral vaccines.
 SQ Sequence 206 AA;
 SQ 17 A; 13 R; 6 N; 10 D; 0 B; 3 C; 6 Q; 19 E; 0 Z; 16 G; 9 H;
 SQ 4 I; 17 L; 10 K; 4 M; 7 F; 15 P; 11 S; 10 T; 7 W; 7 Y; 15 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 CYKLVPVEPDKVEEANKGENTSLHPVSLHGMDDPEREVLEWRFD SRLAFHHVARELHPEYFKNC
 150 160 170 180 190 X X 200

14. US-08-249-182-1 (1-5)

P61515 Sequence of E' protein.

ID P61515 standard; Protein; 206 AA.
 AC P61515;
 DT 08-JUN-1991 (first entry)
 DE Sequence of E' protein.
 KW HIV; LAV; AIDS; diagnosis; vaccine.
 OS HTLV-III_B/H9 cells (ATCC CRL 8543).
 PN EP-187041-A.
 PD 09-JUL-1986.
 PF 23-DEC-1985; 309454.
 PR 24-DEC-1984; US-685272.
 PR 04-DEC-1985; US-805069.
 PA (GETH) GENENTECH INC.
 PI Capon DJ, Lasky LA;
 DR WPI; 86-177602/28.
 DR N-PSDB; N60288
 PT Acquired immune deficiency syndrome polypeptide(s) - obtd. by
 PT molecular cloning etc. and used for diagnosis and in vaccines
 PT against virus disease
 PS Example; fig 2; 125pp; English.
 CC A comparison of N60287 with the cDNA of the HTLV-III genome
 CC revealed one particular clone, designated p7.11 which contained a
 CC DNA sequence encoding this peptide (P60308) sequence. This approx.
 CC 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',
 CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base

pair 5 to the gag region (see NC02557).
SQ Sequence 206 AA;
SQ 18 A; 13 R; 7 N; 10 D; 0 B; 3 C; 6 Q; 18 E; 0 Z; 16 G; 9 H;
SQ 6 I; 17 L; 11 K; 4 M; 7 F; 15 P; 10 S; 9 T; 7 W; 7 Y; 13 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
CYKLVPVEPDKVEEANKGENTSLHPVTLHGMDPPEREVLEWRFD SRLAFHHVARELHPEYFKNC
150 160 170 180 190 X X 200

15. US-08-249-182-1 (1-5)

P60423 Sequence of LAV virus ORF F protein.

ID P60423 standard; Protein; 216 AA.
AC P60423;
DT 20-AUG-1991 (first entry)
DE Sequence of LAV virus ORF F protein.
KW AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.
OS Lymphadenopathy virus.
PN W08602383-A.
PD 24-APR-1986.
PF 18-OCT-1985; E00548.
PR 18-OCT-1984; FR-016013.
PR 16-NOV-1984; GB-029099.
PR 21-JAN-1985; GB-001473.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INSP) INST PASTEUR.
PI Montagnier L, Krust B, Chanaret S, Clavel F, Chermann J-C,
PI Barre-Sinoussi F, Alizon M, Sonigo P, Stewart C, Danos O,
PI Wain-Hobson S.
DR WPI; 86-119166/18.
DR N-PSDB; N60365.
PT Purified glyco:protein and peptide(s) - are recognised by sera contg.
PT antibodies against lymphadenopathy virus and useful in detecting
PT AIDS antibodies or in vaccines
PS Disclosure; Fig 4; 75pp; English.
CC The inventors claim a polypeptide which is recognised by sera of
CC human origin contg. antibodies against the virus of
CC lymphadenopathies (LAV) or acquired immune deficiency syndrome
CC (AIDS). Also claimed are various peptides corresp. to the AA
CC sequences deducible from proteins encoded by LAV DNA, defined by
CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance
CC with a formula given in the specification.
SQ Sequence 216 AA;
SQ 18 A; 12 R; 6 N; 10 D; 0 B; 4 C; 7 Q; 18 E; 0 Z; 17 G; 10 H;
SQ 4 I; 17 L; 12 K; 4 M; 8 F; 16 P; 11 S; 11 T; 8 W; 8 Y; 15 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.92
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
CYKLVPVEPDKVEEANKGENTSLHPVSLHGMDPPEREVLEWRFD SRLAFHHVARELHPQYFKNC
160 170 180 190 200 X X 210

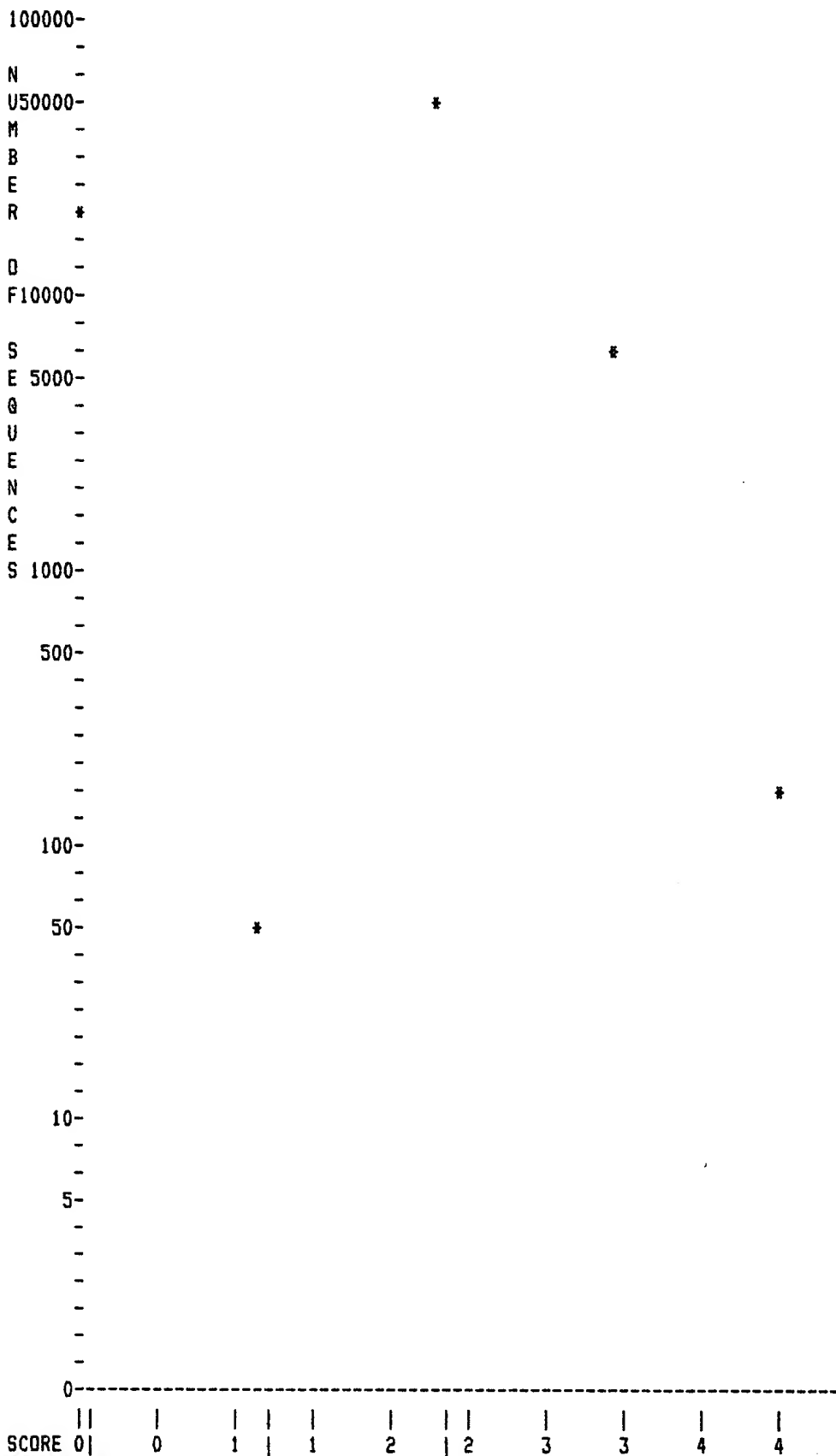
Results file u249_1p.res made by on Thu 22 Sep 94 10:46:43-PDT.

Query sequence being compared:US-08-249-182-1 (1-5)

Number of sequences searched: 70848

Number of scores above cutoff: 3849

Results of the initial comparison of US-08-249-182-1 (1-5) with:
Data bank : PIR 41, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.98

Times:	CPU	Total Elapsed
	00:01:22.94	00:01:32.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	3849

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 3 standard deviations above mean ****						
1. S12146	hypothetical protein E16 - ph	26	4	4	3.05	0
2. J01620	hypothetical 4.3K protein - h	34	4	4	3.05	0
3. PN0630	Acid phosphatase (EC 3.1.3.2)	38	4	4	3.05	0
4. B44255	macrophage mannose receptor,	71	4	4	3.05	0
5. A42329	autotaxin - human (fragments)	114	4	4	3.05	0
6. B45065	putative helicase - Escherich	130	4	4	3.05	0
7. B49205	virulence-associated=vapC - D	135	4	4	3.05	0
8. A40663	acid phosphatase (EC 3.1.3.2)	146	4	4	3.05	0
9. S33783	hemoglobin V - Tokunagayusuri	152	4	4	3.05	0
10. S38404	hypothetical protein - yeast	154	4	4	3.05	0
11. A38148	acid phosphatase (EC 3.1.3.2)	158	4	4	3.05	0
12. D25973	pertussis toxin chain S2 prec	169	4	4	3.05	0
13. XECC3	chloramphenicol O-acetyltrans	203	4	4	3.05	0
14. S03246	nef protein (clone HAT3) - hu	204	4	4	3.05	0
15. S33986	nef protein - human immunodef	206	4	4	3.05	0
16. S25937	nef protein - human immunodef	206	4	4	3.05	0
17. S18691	nef protein - human immunodef	206	4	4	3.05	0
18. S14609	nef protein - human immunodef	206	4	4	3.05	0
19. S03245	nef protein (clone HXB3) - hu	206	4	4	3.05	0
20. S03244	nef protein (clone HXB2) - hu	206	4	4	3.05	0
21. ASLJFV	nef protein - human immunodef	206	4	4	3.05	0
22. ASLJVL	nef protein - human immunodef	206	4	4	3.05	0
23. ASLJ12	nef protein - human immunodef	206	4	4	3.05	0

24. XAEBCP	chloramphenicol O-acetyltrans	213	4	4	3.05	0
25. I44001	nef protein - human immunodef	214	4	4	3.05	0
26. S26084	ribosomal protein S3 - Euglen	218	4	4	3.05	0
27. S34524	ribosomal protein S3 - Euglen	218	4	4	3.05	0
28. S16554	GTP-binding protein rgp1 - ri	226	4	4	3.05	0
29. C25973	pertussis toxin chain S2 prec	226	4	4	3.05	0
30. WEBR2P	pertussis toxin chain S2 prec	226	4	4	3.05	0
31. S23143	acid phosphatase C - human	227	4	4	3.05	0
32. S23142	acid phosphatase A - human	227	4	4	3.05	0
33. A38535	hypothetical protein (MicA re	229	4	4	3.05	0
34. S07475	early light-induced protein,	231	4	4	3.05	0
35. A37147	bacteriophage resistance prot	234	4	4	3.05	0
36. S30700	hypothetical protein o235 - E	235	4	4	3.05	0
37. B37841	hypothetical protein 235 (dap	235	4	4	3.05	0
38. B1AGA6	virB1 protein - Agrobacterium	239	4	4	3.05	0
39. B1AG55	virB1 protein precursor - Agr	239	4	4	3.05	0
40. S15685	kallikrein, glandular - multi	250	4	4	3.05	0

1. US-08-249-182-1 (1-5)

S12146 hypothetical protein E16 - phage D108

ENTRY S12146 #type complete
 TITLE hypothetical protein E16 - phage D108
 ORGANISM #formal_name phage D108
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change 21-Nov-1993
 ACCESSIONS S12146
 REFERENCE S12145
 #authors Pato, M.L.; Banerjee, M.; Wagonner, B.T.
 #journal Nucleic Acids Res. (1990) 18:6458
 #title Sequence of gene E15 of bacteriophage D108 and comparison with phage Mu.
 #cross-references MUID:91057162
 #accession S12146
 ##status preliminary
 ##residues 1-26 ##label PAT
 ##cross-references EMBL:X54298
 SUMMARY #length 26 #molecular-weight 3102 #checksum 6658
 SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      HHVAR
      ||||
MSRTSLIKLIHVARRELQDDDTYRA
    10  X    20

```

2. US-08-249-182-1 (1-5)

JQ1620 hypothetical 4.3K protein - human immunodeficiency

ENTRY JQ1620 #type complete
 TITLE hypothetical 4.3K protein - human immunodeficiency virus type 1
 ORGANISM #formal_name human immunodeficiency virus type 1, HIV-1
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change 21-Nov-1993
 ACCESSIONS JQ1620
 REFERENCE JQ1620
 #authors Smith, J.; Azad, A.; Deacon, N.
 #journal J. Gen. Virol. (1992) 73:1825-1828
 #title Identification of two novel human immunodeficiency virus type

1 splice acceptor sites in infected T cell lines.

#cross-references MUID:92333269
#accession J01620
##status preliminary
##residues 1-34 ##label SMI
SUMMARY #length 34 #molecular-weight 4274 #checksum 4556
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 ||||
MDDPEREVLEWRFDSRLAFHHVARELHPEYFKNC
 10 20 X 30

3. US-08-249-182-1 (1-5)

PN0630 Acid phosphatase (EC 3.1.3.2) 1F, low-molecular-we

ENTRY PN0630 #type fragment
TITLE Acid phosphatase (EC 3.1.3.2) 1F, low-molecular-weight -
Human (fragment)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 27-May-1994; #sequence_revision 27-May-1994; #text_change
27-May-1994
ACCESSIONS PN0630
REFERENCE PN0630
#authors Lazaruk, K.D.A.; Dissing, J.; Sensabaugh, G.F.
#journal Biochem. Biophys. Res. Commun. (1993) 196:440-446
#title Exon structure at the human ACP1 locus supports alternative
splicing model for f and s isozyme generation.
#accession PN0630
##status preliminary
##residues 1-38 ##label LAZ
SUMMARY #length 38 #checksum 7022
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 ||||
WRVDSAATSGYEIGNPPDYRGQSCMKRHGIPMSHVARQ
 10 20 30 X X

4. US-08-249-182-1 (1-5)

B44255 macrophage mannose receptor, MRC1=group I hepatic

ENTRY B44255 #type fragment
TITLE macrophage mannose receptor, MRC1=group I hepatic
glycoprotein receptor homolog (carbohydrate-recognition
domains 1-8) - human (fragment)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 30-Apr-1993; #sequence_revision 30-Apr-1993; #text_change
30-Apr-1993
ACCESSIONS B44255
REFERENCE A44255
#authors Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.
#journal Genomics (1992) 14:721-727
#title Organization of the gene encoding the human macrophage

mannose receptor (MRC1).

#cross-references MUID:93052405
#accession B44255
##status preliminary
##residues 1-71 ##label KIM
##cross-references NCBIP:118421
##note sequence extracted from NCBI backbone
SUMMARY #length 71 #checksum 5821
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
|| ||
FEGSESLWNKDPLTSVSYQINKSALLTWHQARKSCQQAELLSITEIHEQTYLTGLTSSLTSGLWIGLNS
10 20 30 X 40 50 60 70

5. US-08-249-182-1 (1-5)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
TITLE autotaxin - human (fragments)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
08-May-1993
ACCESSIONS A42329
REFERENCE A42329
#authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
Ciocce, V.; Schiffmann, E.; Liotta, L.A.
#journal J. Biol. Chem. (1992) 267:2524-2529
#title Identification, purification, and partial sequence analysis
of autotaxin, a novel motility-stimulating protein.
#cross-references MUID:92129337
#accession A42329
##status preliminary
##molecule_type protein
##residues 1-114 ##label STR
##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
NCBIP:78509; NCBIP:78508; NCBIP:78503
##note sequence extracted from NCBI backbone
SUMMARY #length 114 #checksum 7335
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
|||
TATKSPPFENINLYYDVPWNETIPEEVTPXPNYLQAEVSYPAFKXPXLDVYKWHVAAN
60 70 80 90 100 110 X

6. US-08-249-182-1 (1-5)

B45065 putative helicase - Escherichia coli (fragment)

ENTRY B45065 #type fragment
TITLE putative helicase - Escherichia coli (fragment)
ORGANISM #formal_name Escherichia coli
DATE 10-Jun-1993; #sequence_revision 10-Jun-1993; #text_change

10-Jun-1993
 ACCESSIONS B45065
 REFERENCE A45065
 #authors Huang, S.; Deutscher, M.P.
 #journal J. Biol. Chem. (1992) 267:25609-25613
 #title Sequence and transcriptional analysis of the Escherichia coli
 rnt gene encoding RNase T.
 #cross-references MUID:93094287
 #contents K12 strain UT481
 #accession B45065
 ##status preliminary
 ##molecule_type nucleic acid
 ##residues 1-130 ##label HUA
 ##cross-references NCBIP:120478
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 130 #checksum 1961
 SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 ||||
 MADNPDPSLLPDVFS PATRDWFLRAFKQPTAVQPTWHVHAARSEHALVIAPTGS GKT LA AFLYALDRLFRE
 10 20 30 40 X 50 60 70

 GGEDTREAHKRKTSRILYIS
 80 90

7. US-08-249-182-1 (1-5)

B49205 virulence-associated=vapC - Dichelobacter nodosus

ENTRY B49205 #type complete
 TITLE virulence-associated=vapC - Dichelobacter nodosus
 ORGANISM #formal_name Dichelobacter nodosus
 DATE 19-Dec-1993; #sequence_revision 19-Dec-1993; #text_change
 19-Dec-1993
 ACCESSIONS B49205
 REFERENCE A49205
 #authors Katz, M.E.; Strugnell, R.A.; Rood, J.I.
 #journal Infect. Immun. (1992) 60:4586-4592
 #title Molecular characterization of a genomic region associated
 with virulence in Dichelobacter nodosus.
 #cross-references MUID:93014173
 #accession B49205
 ##status preliminary
 ##molecule_type nucleic acid
 ##residues 1-135 ##label KAT
 ##cross-references NCBIN:116426; NCBIP:116429
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 135 #molecular-weight 15128 #checksum 5461
 SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 || ||
 RYEIGDIGISNITACELAFGVEKSGSAKNKTALT KFLAPLSILPFDKQAIWHYARIRQSLQNRGTPIGALDM
 30 40 50 60 70 80 X 90 100

8. US-08-249-182-1 (1-5)

A40663 acid phosphatase (EC 3.1.3.2) alpha, adipocyte - h

ENTRY A40663 #type fragment
TITLE acid phosphatase (EC 3.1.3.2) alpha, adipocyte - human
(fragment)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 16-Feb-1994; #sequence_revision 16-Feb-1994; #text_change
16-Feb-1994
ACCESSIONS A40663
REFERENCE A40663
#authors Shekels, L.L.; Smith, A.J.; Van Etten, R.L.; Bernlohr, D.A.
#journal Protein Sci. (1992) 1:710-721
#title Identification of the adipocyte acid phosphatase as a
PAD-sensitive tyrosyl phosphatase.
#cross-references MUID:93284125
#contents adipocytes
#accession A40663
##status preliminary
##molecule_type nucleic acid
##residues 1-146 ##label SHE
##cross-references NCBIN:134183; NCBIP:134185
##note sequence extracted from NCBI backbone
KEYWORDS phosphoric monoester hydrolase
SUMMARY #length 146 #checksum 1878
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
IAEAVFRKLVTDQNISENWRVDSAATSGYEIGNPPDYRGQSCMKRHGIPMSHVARQITKEDFATFDYILCMD
10 20 30 40 50 60 X 70 80

ESNLRDLNRKSNRVKTKAKIELLSYDPQKQL
90 100 110

9. US-08-249-182-1 (1-5)

S33783 hemoglobin V - Tokunagayusurika akanusi

ENTRY S33783 #type complete
TITLE hemoglobin V - Tokunagayusurika akanusi
ORGANISM #formal_name Tokunagayusurika akanusi
DATE 02-Dec-1993; #sequence_revision 02-Dec-1993; #text_change
02-Dec-1993
ACCESSIONS S33783
REFERENCE S33783
#authors Fukuda, M.; Takagi, T.; Shikama, K.
#journal Biochim. Biophys. Acta (1993) 1157:185-191
#title Polymorphic hemoglobin from a midge larva (Tokunagayusurika
akanusi) can be divided into two different types.
#accession S33783
##status preliminary
##residues 1-152 ##label FUK
SUMMARY #length 152 #molecular-weight 17197 #checksum 8528
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05

Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                WHVAR
                                ||||
ETLKGHPLDEVKDTANFKLIAGRIFTIFDNCVKNVGNKGFQKVIADMSGPHVARPITHGSYNDLRGVIIYDS
 50      60      70      80      90     100     110

MHL DSTHGAAWNKMMDNFFVVFYECLDGRCSQF
120     130     140     150
```

10. US-08-249-182-1 (1-5)

S38404 hypothetical protein - yeast (*Saccharomyces cerevi*

ENTRY S38404 #type complete
TITLE hypothetical protein - yeast (*Saccharomyces cerevisiae*)
ORGANISM #formal_name *Saccharomyces cerevisiae*
DATE 17-Mar-1994; #sequence_revision 17-Mar-1994; #text_change
17-Mar-1994
ACCESSIONS S38404
REFERENCE S37786
#authors Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger,
F.
#submission submitted to the EMBL Data Library, September 1993
#accession S38404
##status preliminary
##residues 1-154 ##label VAN
##cross-references EMBL:Z26878
SUMMARY #length 154 #molecular-weight 17967 #checksum 6460
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                WHVAR
                                ||||
AIQRNDGGSSENSYGTITGFQDRNRYYPRIQGHDDPCHRIVHYLLIPRIHVARFFHFETLSLVGHMQYI
40      50      60      70      80      90  X  100     110

LAALIARAECTCLGHNTRLLISFWLISQLGTFS
120     130     140
```

11. US-08-249-182-1 (1-5)

A38148 acid phosphatase (EC 3.1.3.2) A - human

ENTRY A38148 #type complete
TITLE acid phosphatase (EC 3.1.3.2) A - human
ALTERNATE_NAMES acid phosphatase B-f; HCPTP-A; phosphotyrosyl protein
phosphatase A
ORGANISM #formal_name *Homo sapiens* #common_name man
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
31-Dec-1993
ACCESSIONS A38148; A39491
REFERENCE A38148
#authors Wo, Y.Y.P.; McCormack, A.L.; Shabanowitz, J.; Hunt, D.F.;
Davis, J.P.; Mitchell, G.L.; Van Etten, R.L.
#journal J. Biol. Chem. (1992) 267:10856-10865
#title Sequencing, cloning, and expression of human red cell-type
acid phosphatase, a cytoplasmic phosphotyrosyl protein
phosphatase.
#cross-references MUID:92268143

```

#contents      red cells
#accession     A38148
##molecule_type mRNA
##residues     1-158 ##label W01
##cross-references NCBIP:103842
##note         sequence extracted from NCBI backbone
REFERENCE      A39491
#authors       Dissing, J.; Johnsen, A.H.; Sensabaugh, G.F.
#journal        J. Biol. Chem. (1991) 266:20619-20625
#title          Human red cell acid phosphatase (ACP1). The amino acid
                 sequence of the two isozymes Bf and Bs encoded by the
                 ACP1*B allele.
#cross-references MUID:92041911
#accession     A39491
##molecule_type protein
##residues     2-158 ##label DIS
KEYWORDS        phosphoric monoester hydrolase
SUMMARY         #length 158 #molecular-weight 18042 #checksum 5808
SEQUENCE

```

```

Initial Score   =      4   Optimized Score =      4   Significance = 3.05
Residue Identity =    80%   Matches         =      4   Mismatches  =      1
Gaps            =      0   Conservative Substitutions =      0

```

```

                                     X  X
                                     WHVAR
                                     ||||
IAEAVFRKLVTDQNISENWRVDSAATSGYEIGNPPDYRGSGCMKRHGIPMSHVARQITKEDFATFDYILCMD
      30          40          50          60          70 X  X  80          90

ESNLRDLNRKSNQVKTCKAKIELLSYDPQKQL
      100         110         120

```

12. US-08-249-182-1 (1-5)

D25973 pertussis toxin chain S2 precursor - Bordetella pa

```

ENTRY          D25973      #type complete
TITLE           pertussis toxin chain S2 precursor - Bordetella parapertussis
ORGANISM        #formal_name Bordetella parapertussis
DATE            04-Mar-1988 #sequence_revision 04-Mar-1988 #text_change
                 18-Jun-1993
ACCESSIONS      D25973
REFERENCE       A25973
#authors        Arico, B.; Rappuoli, R.
#journal         J. Bacteriol. (1987) 169:2847-2853
#title          Bordetella parapertussis and Bordetella bronchiseptica
                 contain transcriptionally silent pertussis toxin genes.
#cross-references MUID:87222217
#accession      D25973
##status        preliminary
##molecule_type DNA
##residues      1-169 ##label ARI
CLASSIFICATION  #superfamily pertussis toxin chain S2
SUMMARY         #length 169 #molecular-weight 18217 #checksum 656
SEQUENCE

```

```

Initial Score   =      4   Optimized Score =      4   Significance = 3.05
Residue Identity =    80%   Matches         =      4   Mismatches  =      1
Gaps            =      0   Conservative Substitutions =      0

```

```

                                     X  X
                                     WHVAR
                                     ||||
MPISRKTLCHLLSVLPLAFLGCHVARASTPGIVIPPQEQITQHGPGPYGRKANKTRALTVAELRGSGDLQEYL
      10          20 X  X  30          40          50          60          70

```

CLASSIFICATION #superfamily AIDS nef protein
SUMMARY #length 204 #molecular-weight 23176 #checksum 8284
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
CFKLVPVEPDKVEEATEGENNSLLHPICLHGMDPEKEVLVWKFD SRLAFHHVAREKHPEYYKDC
140 150 160 170 180 190 X 200

15. US-08-249-182-1 (1-5)

S33986 nef protein - human immunodeficiency virus type 1

ENTRY S33986 #type complete
TITLE nef protein - human immunodeficiency virus type 1
ORGANISM #formal_name human immunodeficiency virus type 1, HIV-1
DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993

ACCESSIONS S33986

REFERENCE S33979

#authors Carlini, F.

#submission submitted to the EMBL Data Library, November 1991

#accession S33986

##status preliminary

##residues 1-206 ##label CAR

##cross-references EMBL:Z11530

SUMMARY #length 206 #molecular-weight 23398 #checksum 2728

SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.05
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
WHVAR
||||
CYKLVPVEPEKLEEANKGENTSLLHPVSLHGMDPGREVLEWRFD SRLAFHHVARELHPEYFKNC
150 160 170 180 190 X X 200

> O <
O| |O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_1s.res made by on Thu 22 Sep 94 10:49:02-PDT.

Query sequence being compared:US-08-249-182-1 (1-5)
Number of sequences searched: 36000
Number of scores above cutoff: 3810

Results of the initial comparison of US-08-249-182-1 (1-5) with:
Data bank : Swiss-Prot 28, all entries

100000-
-
N -
U50000-
M -
B -

13. US-08-249-182-1 (1-5)

XXECC3 chloramphenicol O-acetyltransferase (EC 2.3.1.28)

ENTRY XXECC3 #type complete
 TITLE chloramphenicol O-acetyltransferase (EC 2.3.1.28) III -
 Escherichia coli
 ORGANISM #formal_name Escherichia coli
 DATE 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change
 30-Jun-1993
 ACCESSIONS A00567
 REFERENCE A00567
 #authors Packman, L.C.; Kaye, N.M.C.; Fitton, J.E.
 #citation unpublished results, cited by Shaw, W.V., CRC Crit. Rev.
 Biochem. 14, 1-46, 1983
 #accession A00567
 ##molecule_type protein
 ##residues 1-203 ##label PAC
 CLASSIFICATION #superfamily chloramphenicol acetyltransferase
 KEYWORDS acyltransferase; antibiotic resistance
 SUMMARY #length 203 #molecular-weight 23824 #checksum 9862
 SEQUENCE

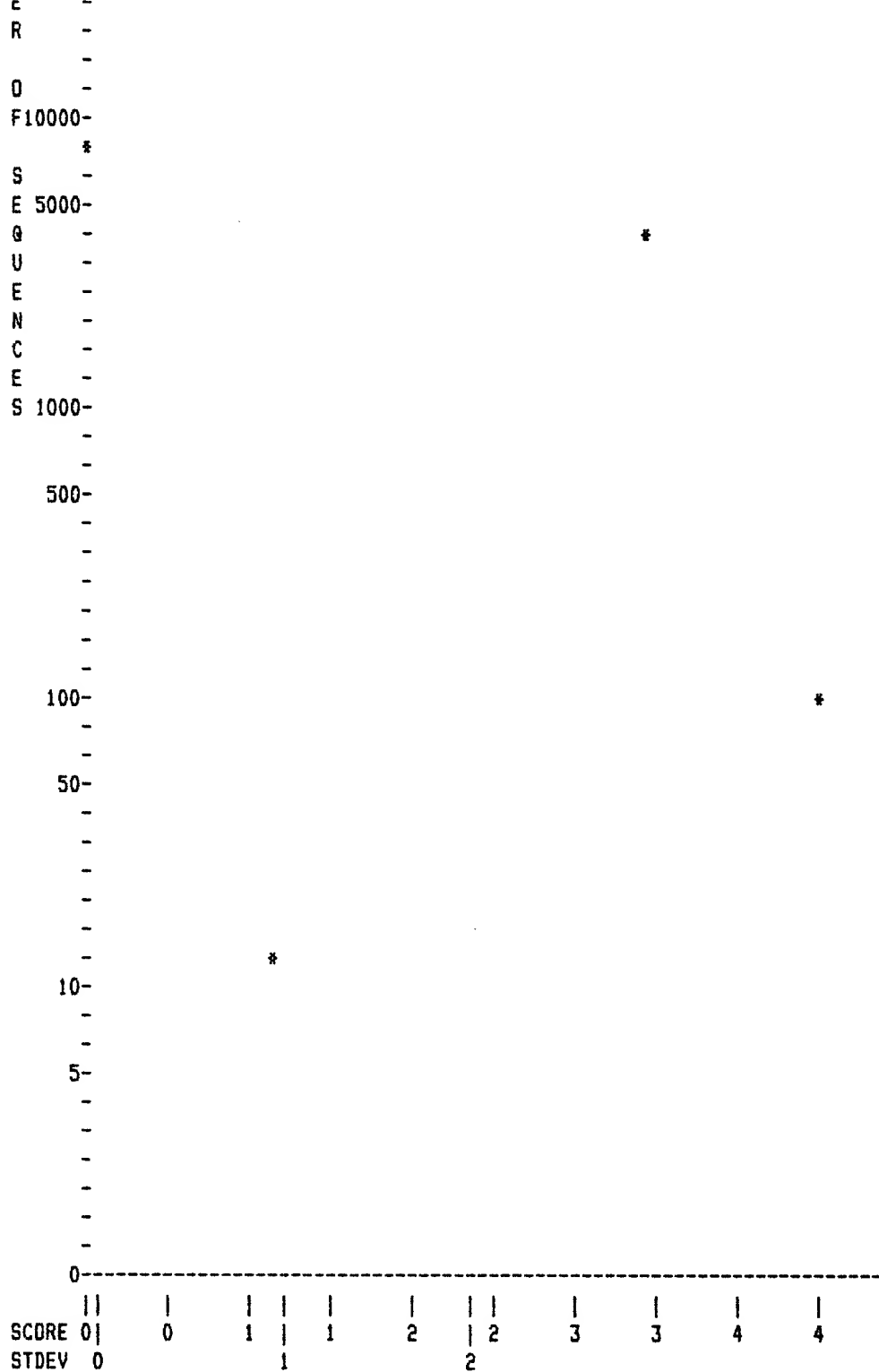
Initial Score = 4 Optimized Score = 4 Significance = 3.05
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 IIII
 INPLAPWVNFDSFDLNVANFDNMAKYQEGDRLLPLVLSVGVHHA VCDGFHVARFINRLQELCNSKLK
 140 150 160 170 180 X 190 200

14. US-08-249-182-1 (1-5)

S03246 nef protein (clone HAT3) - human immunodeficiency

ENTRY S03246 #type complete
 TITLE nef protein (clone HAT3) - human immunodeficiency virus type
 1
 ALTERNATE_NAMES 3'-orf protein
 ORGANISM #formal_name human immunodeficiency virus type 1, HIV-1
 DATE 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change
 30-Sep-1993
 ACCESSIONS S03246
 REFERENCE S03244
 #authors Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy,
 E.P.; Livak, K.J.; Petteway Jr., S.R.; Pearson, M.L.;
 Haseltine, W.A.; Arya, S.K.; Wong-Staal, F.
 #journal Nucleic Acids Res. (1985) 13:8219-8229
 #title Polymorphism of the 3' open reading frame of the virus
 associated with the acquired immune deficiency syndrome,
 human T-lymphotropic virus type III.
 #cross-references MUID:86067228
 #accession S03246
 ##molecule_type DNA
 ##residues 1-204 ##label RAT
 ##cross-references EMBL:X03190
 ##note the authors translated the codon AGT for residue 11 as
 Gly
 GENETICS
 #gene nef; 3'-orf; orf-F



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.89

Times: CPU Total Elapsed
00:00:48.92 00:00:54.00

Number of residues: 12496420
Number of sequences searched: 36000
Number of scores above cutoff: 3810

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Score	Sig.	Frame
**** 3 standard deviations above mean ****							
1. E16_BPD10	PROTEIN E16 (FRAGMENT).	26	4	4	3.38	0	
2. YBBD_ECOLI	HYPOTHETICAL 9.2 KD PROTEIN I	78	4	4	3.38	0	
3. CYB_SCAPL	CYTOCHROME B (EC 1.10.2.2) (F	98	4	4	3.38	0	
4. RHLF_ECOLI	PUTATIVE ATP-DEPENDENT RNA HE	130	4	4	3.38	0	
5. Y145_ADE07	HYPOTHETICAL 14.5 KD EARLY PR	133	4	4	3.38	0	
6. PPAF_HUMAN	RED CELL ACID PHOSPHATASE 1,	157	4	4	3.38	0	
7. NEF_HVIPV	NEGATIVE FACTOR (F-PROTEIN) (206	4	4	3.38	0	
8. NEF_HV1BR	NEGATIVE FACTOR (F-PROTEIN) (206	4	4	3.38	0	
9. NEF_HV112	NEGATIVE FACTOR (F-PROTEIN) (206	4	4	3.38	0	
10. NEF_HV1RH	NEGATIVE FACTOR (F-PROTEIN) (208	4	4	3.38	0	
11. CAT3_ECOLI	CHLORAMPHENICOL ACETYLTRANSFE	213	4	4	3.38	0	
12. NEF_HV1JR	NEGATIVE FACTOR (F-PROTEIN) (216	4	4	3.38	0	
13. RS3_EUGGR	30S RIBOSOMAL PROTEIN S3.	218	4	4	3.38	0	
14. TOX2_BORPE	PERTUSSIS TOXIN SUBUNIT 2 (S2	226	4	4	3.38	0	
15. RGP1_DRYSA	GTP-BINDING REGULATORY PROTEI	226	4	4	3.38	0	
16. YGGH_ECOLI	HYPOTHETICAL 25.9 KD PROTEIN	230	4	4	3.38	0	
17. ELI5_HORVU	HIGH MOLECULAR MASS EARLY LIG	231	4	4	3.38	0	
18. YIGA_ECOLI	HYPOTHETICAL 26.7 KD PROTEIN	235	4	4	3.38	0	
19. VIB1_AGRT9	VIRB1 PROTEIN PRECURSOR.	239	4	4	3.38	0	
20. NGFA_MOUSE	7S NERVE GROWTH FACTOR ALPHA	256	4	4	3.38	0	
21. PRIA_LENED	PRIA PROTEIN PRECURSOR.	258	4	4	3.38	0	
22. NGFG_MOUSE	7S NERVE GROWTH FACTOR GAMMA	261	4	4	3.38	0	
23. KAGB_MOUSE	GLANDULAR KALLIKREIN MGK-11 P	261	4	4	3.38	0	
24. KAG1_MOUSE	GLANDULAR KALLIKREIN MGK-1 PR	261	4	4	3.38	0	
25. EGBC_MOUSE	EPIDERMAL GROWTH FACTOR-BINDI	261	4	4	3.38	0	
26. KAGR_PRANA	GLANDULAR KALLIKREIN PRECURSO	263	4	4	3.38	0	
27. YIHW_ECOLI	HYPOTHETICAL 29.5 KD PROTEIN	269	4	4	3.38	0	
28. TRY4_ANOGA	TRYPSIN 4 PRECURSOR (EC 3.4.2	275	4	4	3.38	0	
29. ETB_STAAV	EXFOLIATIVE TOXIN B PRECURSOR	277	4	4	3.38	0	
30. YIJD_ECOLI	HYPOTHETICAL 32.1 KD PROTEIN	283	4	4	3.38	0	
31. RP32_CITFR	RNA POLYMERASE SIGMA-32 FACTO	284	4	4	3.38	0	
32. YOHI_ECOLI	HYPOTHETICAL 35.2 KD PROTEIN	315	4	4	3.38	0	
33. PSBD_WHEAT	OXYGEN-EVOLVING ENHANCER PROT	325	4	4	3.38	0	
34. UL95_EBV	HYPOTHETICAL PROTEIN BGLF3.	332	4	4	3.38	0	
35. IPNS_STRLP	ISOPENICILLIN N SYNTHETASE (I	333	4	4	3.38	0	
36. GVPN_HALSA	GVPN PROTEIN.	345	4	4	3.38	0	
37. GVPN_HALME	GVPN PROTEIN.	347	4	4	3.38	0	
38. GVPN_HALHA	GVPN PROTEIN.	347	4	4	3.38	0	
39. CYB_ACITR	CYTOCHROME B (EC 1.10.2.2).	380	4	4	3.38	0	
40. RL3A_ARATH	60S RIBOSOMAL PROTEIN L3.	388	4	4	3.38	0	

1. US-08-249-182-1 (1-5)

E16_BPD10 PROTEIN E16 (FRAGMENT).

ID E16_BPD10 STANDARD; PRT; 26 AA.
 AC P24796;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE PROTEIN E16 (FRAGMENT).
 GN E16.
 OS BACTERIOPHAGE D108.
 OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91057162
 RA PATO M.L., BANERJEE M., WAGONNER B.T.;
 RL NUCLEIC ACIDS RES. 18:6458-6458(1990).
 DR EMBL; X54298; BD108E15.
 DR PIR; S12146; S12146.
 FT NON_TER 26 26
 SQ SEQUENCE 26 AA; 3102 MW; 2831 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 MSRTSLIKLIHVARRELQLDDDTYRA
 10 X 20

2. US-08-249-182-1 (1-5)

YBBD_ECOLI HYPOTHETICAL 9.2 KD PROTEIN IN RHSD 3'REGION (ORFD

ID YBBD_ECOLI STANDARD; PRT; 78 AA.
 AC P33669;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 9.2 KD PROTEIN IN RHSD 3'REGION (ORFD3).
 GN YBBD.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RM 92115567
 RA SADOSKY A.B., GRAY J.A., HILL C.W.;
 RL NUCLEIC ACIDS RES. 19:7177-7183(1991).
 DR EMBL; L19084; ECORFD23U.
 DR ECGENE; EG11770; YBBD.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 78 AA; 9157 MW; 33115 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 FIFATIIIVVLCVITYLYLYKDESLVSKHYINYMAIPENDGVFTWLPDFFPHVARGYINIHKCR
 20 30 40 50 60 X 70

3. US-08-249-182-1 (1-5)

CYB_SCAPL CYTOCHROME B (EC 1.10.2.2) (FRAGMENT).

ID CYB_SCAPL STANDARD; PRT; 98 AA.
 AC P29672;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B (EC 1.10.2.2) (FRAGMENT).
 GN COB OR CYTB.
 OS SCAPHIRHYNCHUS PLATORYNCHUS (STURGEON).
 OG MITOCHONDRION.
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
 OC OSTEICHTHYES; ACTINOPTERYGII; ACIPENSERIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92130804
 RA NORMARK B.B., MCCUNE A.R., HARRISON R.G.;
 RL MOL. BIOL. EVOL. 8:819-834(1991).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -!- CATALYTIC ACTIVITY: $QH(2) + 2 \text{ FERRICYTOCHROME C} = Q +$
 CC $2 \text{ FERROCYTOCHROME C.}$
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 DR EMBL; M64921; MISPCDC.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME.
 DR PROSITE; PS00193; CYTOCHROME_B_QO.
 KW ELECTRON TRANSPORT; MITOCHONDRION; RESPIRATORY CHAIN; TRANSMEMBRANE;
 KW HEME.
 FT NON_TER 1 1
 FT METAL 48 48 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 62 62 IRON 2 (HEME B566 AXIAL LIGAND).
 FT NON_TER 98 98
 SQ SEQUENCE 98 AA; 11031 MW; 56757 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 LTGLFLAMHYTADISTAFSSVAHICRDVNYGWLIRNVHANGASFFFCILYLVHVARGMYYGSYLQKETWNIGV
 20 30 40 50 60 X 70 80

VLLLLTMMTAFVGVVL
 90

4. US-08-249-182-1 (1-5)

RHLF_ECOLI PUTATIVE ATP-DEPENDENT RNA HELICASE RHLF (FRAGMENT

ID RHLF_ECOLI STANDARD; PRT; 130 AA.
 AC P30015;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PUTATIVE ATP-DEPENDENT RNA HELICASE RHLF (FRAGMENT).
 GN RHLF.

US ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RM 93094287
 RA HUANG S., DEUTSCHER M.P.;
 RL J. BIOL. CHEM. 267:25609-25613(1992).
 CC -!- SIMILARITY: TO OTHER RNA HELICASES.
 DR EMBL; L01622; ECRNTASET.
 DR ECGENE; EG11548; RHLF.
 DR PIR; B45065; B45065.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE.
 KW HELICASE; ATP-BINDING; RNA-BINDING.
 FT NP_BIND 51 58 ATP (BY SIMILARITY).
 FT NON_TER 130 130
 SQ SEQUENCE 130 AA; 14560 MW; 77456 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                X  X
                WHVAR
                ||||
MADNPDPSLLPDVFS PATRDWFLRAFKOPTAVGPOTWHVAARSEHALVIAPTGSCKTLAAFLYALDRLFRE
      10      20      30      40 X      50      60      70

GGEDTREAHKRKTSRILYIS
      80      90
  
```

5. US-08-249-182-1 (1-5)
 Y145_ADE07 HYPOTHETICAL 14.5 KD EARLY PROTEIN.

ID Y145_ADE07 STANDARD; PRT; 133 AA.
 AC P05667;
 DT 01-NOV-1988 (REL. 09, CREATED)
 DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 14.5 KD EARLY PROTEIN.
 OS HUMAN ADENOVIRUS TYPE 7.
 OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GOMEN;
 RM 83183660
 RA ENGLER J.A., HOPPE M.S., VAN BREE M.P.;
 RL GENE 21:145-159(1983).
 DR EMBL; X03000; AD7001.
 KW HYPOTHETICAL PROTEIN; EARLY PROTEIN.
 SQ SEQUENCE 133 AA; 14557 MW; 80147 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                X  X
                WHVAR
                ||||
MRDDAAVDILDLTPLGESYRPRELEPEREFNRINLGIVDGGLPKDFLHVARVVLVGD LGHELLDLFLLEISA
      10      20      30      40      X 50      60      70

ARSLDGGREVVGDAPNELRESIHARLVPD
      80      90     100
  
```

6. US-08-249-182-1 (1-5)

PPAF_HUMAN RED CELL ACID PHOSPHATASE 1, ISOZYME F (EC 3.1.3.2)

ID PPAF_HUMAN STANDARD; PRT; 157 AA.
 AC P24666;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE RED CELL ACID PHOSPHATASE 1, ISOZYME F (EC 3.1.3.2) (ACP1).
 GN ACP1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP B ALLELE, SEQUENCE.
 RM 92041911
 RA DISSING J., JOHNSEN A.H., SENSABAUGH G.F.;
 RL J. BIOL. CHEM. 266:20619-20625(1991).
 RN [2]
 RP A AND C ALLELES, SEQUENCE.
 RM 92329495
 RA DISSING J., JOHNSEN A.H.;
 RL BIOCHIM. BIOPHYS. ACTA 1121:261-268(1992).
 CC -!- FUNCTION: HAS BEEN SHOWN TO HYDROLYZE FMN AND PROTEIN TYROSINE
 CC PHOSPHATE.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- HUMAN ACP1 IS GENETICALLY POLYMORPHIC, AND THREE COMMON ALLELES
 CC SEGREGATING AT THE ACP1 LOCUS GIVE RISE TO SIX PHENOTYPES. EACH
 CC ALLELE APPEARS TO ENCODE TWO ELECTROPHORETICALLY DIFFERENT
 CC ISOZYMES, F AND S, WHICH ARE PRODUCED IN ALLELE-SPECIFIC RATIOS.
 CC -!- THE SEQUENCE SHOWN IS THAT OF ALLELES B AND C.
 CC -!- ALTERNATIVE SPLICING: THE F AND S ISOZYMES PROBABLY RESULT FROM
 CC ALTERNATIVE SPLICING OF THE PRIMARY RNA TRANSCRIPT. THE RATIO OF
 CC F TO S IS 2:1 IN ALLELE A, 4:1 IN ALLELE B, AND 1:4 IN ALLELE C.
 CC -!- THIS PHOSPHATASE IS INHIBITED BY SULFHYDRYL REAGENTS.
 CC -!- SIMILARITY: TO BOVINE LOW MOLECULAR WEIGHT CYTOSOLIC ACID
 CC PHOSPHATASE.
 DR PIR; A39491; A39491.
 DR MIM; 171500; TENTH EDITION.
 KW HYDROLASE; ACETYLATION; ALTERNATIVE SPLICING.
 FT MOD_RES 1 1 ACETYLATION.
 FT ACT_SITE 12 12 PROBABLE.
 FT ACT_SITE 17 17 PROBABLE.
 FT VARIANT 105 105 Q -> R (IN ALLELE A).
 SQ SEQUENCE 157 AA; 17996 MW; 118642 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 |||

IAEAVFRKLVWDQNISENWRVDSAATSGYEIGNPPDYRGQSCMKRHGIPMSHVARQITKEDFATFDYILCMD
 30 40 50 60 70 X 80 90

ESNLRDLNRKSNQVKTKAKIELLGSYDPQKQL
 100 110 120

7. US-08-249-182-1 (1-5)

NEF_HVIPV NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF)

ID NEF_HV1PV STANDARD; PRT; 206 AA.
AC P03405;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF).
GN NEF.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (PV22 ISOLATE) (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 85111157
RA MUESING M.A., SMITH D.H., CABRADILLA C.D., BENTON C.V., LASKY L.A.,
RA CAPON D.J.;
RL NATURE 313:450-458(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, FUNCTION.
RM 88039140
RA GUY B., KIENV M.-P., RIVIERE Y., LE PEUCH C., DOTT K., GIRARD M.,
RA MONTAGNIER L., LECOCQ J.-P.;
RL NATURE 330:266-269(1987).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; X01762; REHTLV3.
DR HIV; K02083; NEF\$PV22.
DR PIR; A04007; ASLJVL.
KW AIDS; MYRISTYLATION; GTP-BINDING.
FT LIPID 2 2 MYRISTATE.
SQ SEQUENCE 206 AA; 23352 MW; 219653 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
HHVAR
||||
CYKLVPVEPDKVEEANKGENTSLHPVSLHGMDDPEREVLWRFDSRLAFHHVARELHPEYFKNC
150 160 170 180 190 X X 200

8. US-08-249-182-1 (1-5)

NEF_HV1BR NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF)

ID NEF_HV1BR STANDARD; PRT; 206 AA.
AC P03406;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF).
GN NEF.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (BRU ISOLATE) (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 85099333
RA WAIN-HOBSON S., SONIGO P., DANOS O., COLE S., ALIZON M.;
RL CELL 40:9-17(1985).
RN [2]
RP CLONE PNL4-3, SEQUENCE FROM N.A.
RA BUCKLER C.E., BUCKLER-WHITE A.J., WILLEY R.L., MCCOY J.;
RL SUBMITTED (JUN-1988) TO THE HIV DATA BANK.
RN [3]

RF POST-TRANSLATIONAL MODIFICATIONS; FUNCTION.
 RM 88039140
 RA GUY B., KIENY M.-P., RIVIERE Y., LE PEUCH C., DOTT K., GIRARD M.,
 RA MONTAGNIER L., LECOCQ J.-P.;
 RL NATURE 330:266-269(1987).
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
 CC ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
 DR EMBL; K02013; HIVBRUCG.
 DR EMBL; M19921; REHIVNL4.
 DR HIV; K02013; NEF\$BRU.
 DR HIV; M19921; NEF\$NL43.
 DR PIR; A04008; ASLJFV.
 KW AIDS; MYRISTYLATION; GTP-BINDING; PHOSPHORYLATION.
 FT LIPID 2 2 MYRISTATE.
 FT MOD_RES 15 15 PHOSPHORYLATION (BY PKC).
 FT VARIANT 11 11 V -> I (IN CLONE PNL4-3).
 FT VARIANT 15 15 T -> A (IN CLONE PNL4-3).
 FT VARIANT 33 33 A -> V (IN CLONE PNL4-3).
 FT VARIANT 51 51 T -> N (IN CLONE PNL4-3).
 SQ SEQUENCE 206 AA; 23342 MW; 221722 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 WHVAR
 ||||
 CYKLVPVEPKVVEANKGENTSLHPVSLHGMDPEREVLWRFDSRLAFHHVARELHPEYFKNC
 150 160 170 180 190 X X 200

9. US-08-249-182-1 (1-5)

NEF_HV112 NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF)

ID NEF_HV112 STANDARD; PRT; 206 AA.
 AC P04324;
 DT 20-MAR-1987 (REL. 04, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF).
 GN NEF.
 OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (CLONE 12) (HIV-1).
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
 OC LENTIVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 86177573
 RA ARYA S.K., GALLO R.C.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:2209-2213(1986).
 RN [2]
 RP POST-TRANSLATIONAL MODIFICATIONS; FUNCTION.
 RM 88039140
 RA GUY B., KIENY M.-P., RIVIERE Y., LE PEUCH C., DOTT K., GIRARD M.,
 RA MONTAGNIER L., LECOCQ J.-P.;
 RL NATURE 330:266-269(1987).
 CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
 CC ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
 DR EMBL; M11840; HIVDSM.
 DR HIV; M11840; NEF\$PCV12.
 DR PIR; A04006; ASLJ12.
 KW AIDS; MYRISTYLATION; GTP-BINDING.
 FT LIPID 2 2 MYRISTATE.
 SQ SEQUENCE 206 AA; 23366 MW; 218839 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38

Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                WHVAR
                                IIII
CYKLVPVEPEKLEEANKGENTSLHPVSLHGMDDPEREVLEWRFD SRLAFHHVARELHPEYFKNC
150      160      170      180      190 X  X 200
```

10. US-08-249-182-1 (1-5)

NEF_HV1RH NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF)

ID NEF_HV1RH STANDARD; PRT; 208 AA.
AC P05858;
DT 01-NOV-1988 (REL. 09, CREATED)
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF).
GN NEF.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (RF/HAT ISOLATE) (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA STARCICH B.R., HAHN B.H., SHAW G.M., MCNEELY P.D., MODROW S.,
RA WOLF H., PARKS E.S., PARKS W.P., JOSEPHS S.F., GALLO R.C.,
RA WONG-STAAI F.;
RL SUBMITTED (XXX-1987) TO THE HIV DATA BANK.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; M17451; HIVRF.
DR HIV; M17451; NEF*RF.
KW AIDS; MYRISTYLATION; GTP-BINDING.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 208 AA; 23532 MW; 229900 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                WHVAR
                                IIII
CFKLVPVEPKVEEATEGENNSLLHPICLHGMDDPEKEVLVWKFD SRLAFHHVAREKHPEYYKDC
150      160      170      180      190 X 200
```

11. US-08-249-182-1 (1-5)

CAT3_ECOLI CHLORAMPHENICOL ACETYLTRANSFERASE III (EC 2.3.1.28)

ID CAT3_ECOLI STANDARD; PRT; 213 AA.
AC P00484;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE CHLORAMPHENICOL ACETYLTRANSFERASE III (EC 2.3.1.28).
GN CAT3.
OS ESCHERICHIA COLI.
OG PLASMID R387.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 88339790
RA MURRAY I.A., HAWKINS A.R., KEYTE J.W., SHAW W.V.;

RL BIOCHEM. J. 252:173-179(1988).
 RN [2]
 RP PRELIMINARY SEQUENCE.
 RM 83181524
 RA PACKMAN L.C., KAYE N.M.C., FITTON J.E.;
 RL UNPUBLISHED RESULTS, CITED BY:
 RL SHAW W.V.;
 RL CRC CRIT. REV. BIOCHEM. 14:1-46(1983).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RM 88247977
 RA LESLIE A.G.W., MOODY P.C.E., SHAW W.V.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 85:4133-4137(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RM 90250768
 RA LESLIE A.G.W.;
 RL J. MOL. BIOL. 213:167-186(1990).
 CC -!- FUNCTION: THIS ENZYME IS AN EFFECTOR OF CHLORAMPHENICOL RESISTANCE
 CC IN BACTERIA.
 CC -!- CATALYTIC ACTIVITY: ACETYL-COA + CHLORAMPHENICOL = COA +
 CC CHLORAMPHENICOL 3-ACETATE.
 CC -!- SUBUNIT: HOMOTRIMER.
 DR EMBL; X07848; SFCAT3.
 DR PIR; A00567; XECC3.
 DR PIR; S00602; XEBCF.
 DR PDB; 1CLA; 15-OCT-92.
 DR PDB; 2CLA; 15-JUL-90.
 DR PDB; 3CLA; 15-JUL-92.
 DR PDB; 4CLA; 15-JUL-92.
 DR PROSITE; PS00100; CAT.
 KW ANTIBIOTIC RESISTANCE; TRANSFERASE; ACYLTRANSFERASE; PLASMID;
 KW 3D-STRUCTURE.
 FT ACT_SITE 189 189
 FT STRAND 3 5
 FT TURN 9 10
 FT TURN 12 13
 FT HELIX 14 21
 FT TURN 22 23
 FT STRAND 27 35
 FT HELIX 37 44
 FT TURN 45 45
 FT HELIX 50 62
 FT TURN 63 64
 FT HELIX 66 69
 FT STRAND 70 72
 FT STRAND 77 80
 FT STRAND 84 91
 FT TURN 92 95
 FT STRAND 96 101
 FT HELIX 108 121
 FT TURN 123 124
 FT TURN 130 131
 FT STRAND 138 144
 FT TURN 145 146
 FT TURN 160 161
 FT STRAND 166 170
 FT STRAND 173 175
 FT TURN 176 177
 FT STRAND 178 188
 FT TURN 189 191
 FT HELIX 194 208
 FT TURN 209 209
 SQ SEQUENCE 213 AA; 24993 MW; 238808 CN;

Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

X X
WHVAR
||||
PWNVFDSENLNVANFTDYFAPITMAKYGGEGDRLLPLSVGVHHA VCDGFHVARFINRLQELCNSKLK
150      160      170      180      190 X 200      210
```

12. US-08-249-182-1 (1-5)

NEF_HV1JR NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF)

ID NEF_HV1JR STANDARD; PRT; 216 AA.
AC P20867;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE NEGATIVE FACTOR (F-PROTEIN) (27 KD PROTEIN) (3'ORF).
GN NEF.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (JRCSF ISOLATE) (HIV-1).
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC LENTIVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RA KOYANAGI S., CHEN I.S.Y.;
RL SUBMITTED (DEC-1988) TO THE HIV DATA BANK.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES, IT SEEM TO DOWN REGULATE THE CD4(T4) ANTIGEN.
DR EMBL; M38429; HIVJRCSF.
DR HIV; M38429; NEF\$JRCSF.
KW AIDS; MYRISTYLATION; GTP-BINDING.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
SQ SEQUENCE 216 AA; 24567 MW; 243460 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

X X
WHVAR
||||
CFKLV PVDPEKVEEANE GENNCLLHPMSQHGMDDPEKEVLVWKFD SKLALHHVARELHPEYYKDC
160      170      180      190      200 X X 210
```

13. US-08-249-182-1 (1-5)

RS3_EUGGR 30S RIBOSOMAL PROTEIN S3.

ID RS3_EUGGR STANDARD; PRT; 218 AA.
AC P19169;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE 30S RIBOSOMAL PROTEIN S3.
GN RPS3.
OS EUGLENA GRACILIS.
OG CHLOROPLAST.
OC EUKARYOTA; PLANTA; PHYCOPHYTA; EUGLENOPHYTA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z;
RM 89063445
RA CHRISTOPHER D.A., CUSHMAN J.C., PRICE C.A., HALICK R.B.;
RL CURR. GENET. 14:275-286(1988).
CC -!- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.

DR EMBL; Z11874; CHEGZ.
 DR EMBL; M37463; CHEGRIBP.
 DR PIR; S26084; S26084.
 DR PIR; S34524; S34524.
 DR PROSITE; PS00548; RIBOSOMAL_S3_1.
 DR PROSITE; PS00734; RIBOSOMAL_S3_2.
 KW RIBOSOMAL PROTEIN; CHLOROPLAST.
 SQ SEQUENCE 218 AA; 25267 MW; 242884 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                WHVAR
                                |||
VERRHYASFVKEDIVIRNFMNKELLETLSLIKIERIYEFSEQRNNTIVYIHVARPERVIGRDGGLSRIRD
      30      40      50      60      70  X  80      90

ILIDRMNVLLGKTPRIITCKVVGVTSPNLDARL
      100     110     120

```

14. US-08-249-182-1 (1-5)

TOX2_BORPE PERTUSSIS TOXIN SUBUNIT 2 (S2) PRECURSOR (ISLET-AC

ID TOX2_BORPE STANDARD; PRT; 226 AA.
 AC P04978;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
 DE PERTUSSIS TOXIN SUBUNIT 2 (S2) PRECURSOR (ISLET-ACTIVATING PROTEIN)
 DE (IAP).
 OS BORDETELLA PERTUSSIS.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC ALCALIGENACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BP165;
 RM 86259651
 RA NICOSIA A., PERUGINI M., FRANZINI C., CASAGLI M.C., BORRI M.G.,
 RA ANTONI G., ALMONI M., NERI P., RATTI G., RAPPUOLI R.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:4631-4635(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RM 86208173
 RA LOCHT C., KEITH J.M.;
 RL SCIENCE 232:1258-1264(1986).
 CC -!- FUNCTION: B BINDS TO RECEPTORS ON THE EUKARYOTIC CELL SURFACE AND
 CC FACILITATES THE TRANSLOCATION OF THE TOXIC SUBUNIT ACROSS THE CELL
 CC MEMBRANE.
 CC -!- SUBUNIT: PERTUSSIS TOXIN CONTAINS FIVE DIFFERENT CHAINS, S1-S5.
 CC THEY ARE ORGANIZED INTO 2 FUNCTIONAL SUBUNITS: A, COMPOSED OF S1
 CC (WHICH IS TOXIC) AND B, CONTAINING S2, S3, S5, AND TWO COPIES OF
 CC S4 (B BINDS TO THE MEMBRANE RECEPTORS). DIMERS OF S2-S4 AND S3-S4
 CC ARE HELD TOGETHER BY S5.
 DR EMBL; M14378; BPTOXS.
 DR EMBL; M13223; BPTOX.
 DR PIR; B24144; WEBR2P.
 KW MEMBRANE; TOXIN; SIGNAL; WHOOPING COUGH.
 FT SIGNAL 1 27
 FT CHAIN 28 226 TOXIN SUBUNIT 2.
 FT SIMILAR 28 226 67% WITH S3 SUBUNIT.
 SQ SEQUENCE 226 AA; 24829 MW; 287397 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38

Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 ||||
MPIDRKTLCHLLSVLPLALLGSHVARASTPGIVIPPQEQITQHGSPYGRCAKTRALTVAELRGSGDLQEYL
 10 20 X X 30 40 50 60 70

RHVT

15. US-08-249-182-1 (1-5)

RGPI_ORYSA GTP-BINDING REGULATORY PROTEIN RGP1.

ID RGPI_ORYSA STANDARD; PRT; 226 AA.
AC P25766;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE GTP-BINDING REGULATORY PROTEIN RGP1.
GN RGP1.
OS DRYZA SATIVA (RICE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
OC CYPERALES; GRAMINEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. GINBOZU;
RM 91360069
RA SAND H., YOUSSEFIAN S.;
RL MOL. GEN. GENET. 228:227-232(1991).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN PLANT GROWTH AND
CC DEVELOPMENT.
CC -!- DEVELOPMENTAL STAGE: DURING SEEDLING RGP1 EXPRESSION WAS FIRST
CC OBSERVED 14 DAYS AFTER GERMINATION, REACHING A MAXIMUM LEVEL
CC BETWEEN 28 AND 42 DAYS, AND GRADUALLY DECREASED THEREAFTER UNTIL
CC 63 DAYS WHEN IT ATTAINED THE SAME LEVEL OF EXPRESSION AS IN 14-DAY
CC OLD SEEDLINGS.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO YPT1 SUB-FAMILY.
DR EMBL; X59276; OSRGP1.
DR PIR; S16554; S16554.
KW GTP-BINDING; LIPOPROTEIN; PRENYLATION.
FT NP_BIND 25 32 GTP (BY SIMILARITY).
FT NP_BIND 73 77 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
FT DOMAIN 47 55 EFFECTOR REGION (POTENTIAL).
FT LIPID 223 223 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 224 224 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 226 AA; 24852 MW; 266731 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.38
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

 X X
 WHVAR
 ||||
TRLTHIDARTVKAQTRDTAGQERYRAVTSAYYRGAVGAMLVYDITKRQSFQDHVARWLEELRGHADKNIVIML
 60 70 80 90 100 110 120

IGNKSDLGTLRVVPTEDAKEFAERENLFFMETS
 130 140 150 160

> 0 <
0| 0 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences

Release 5.4

Seq. 2

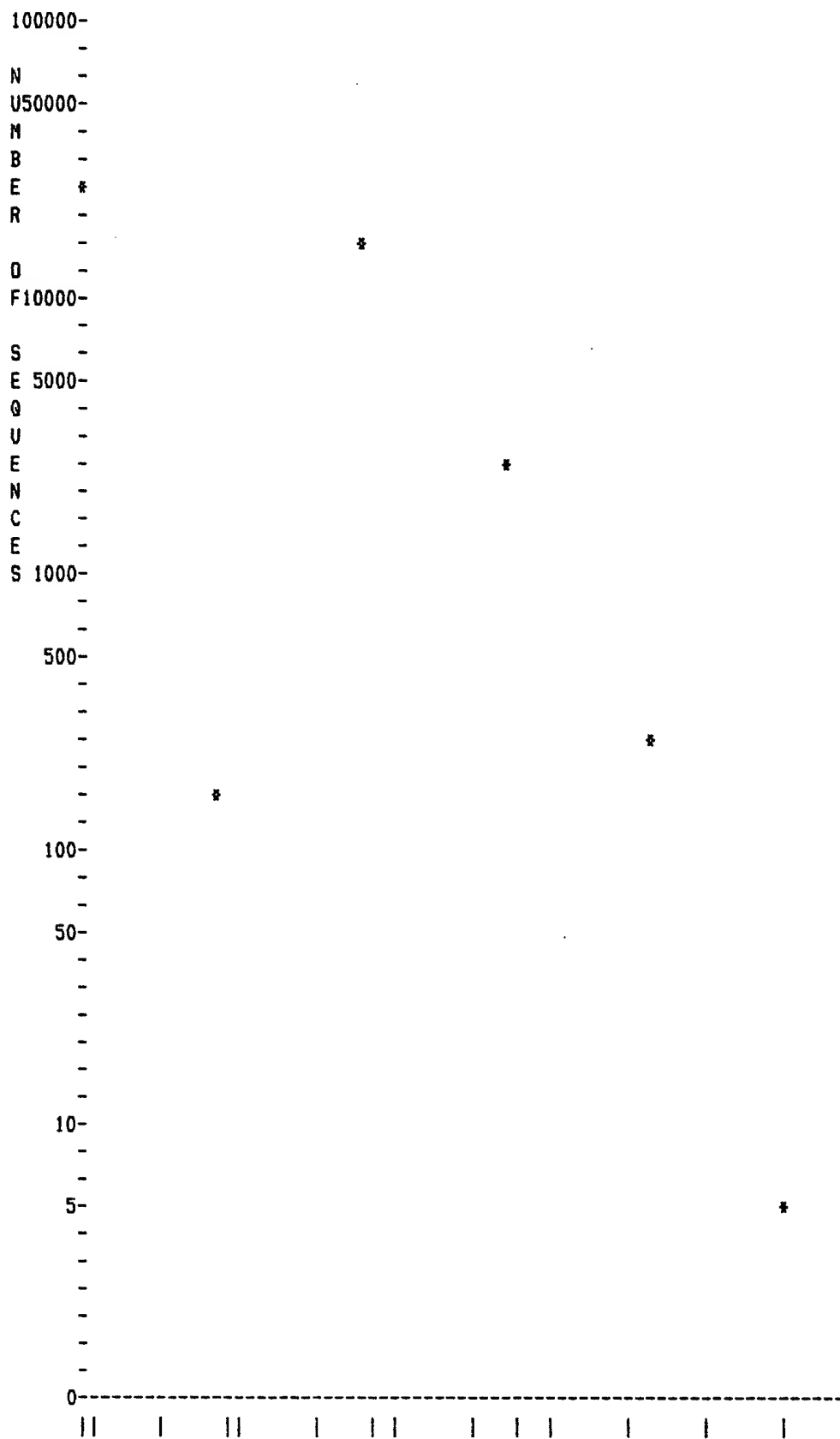
Results file u249_2a.res made by on Thu 22 Sep 94 10:26:37-PDT.

Query sequence being compared: US-08-249-182-2 (1-6)

Number of sequences searched: 42145

Number of scores above cutoff: 3792

Results of the initial comparison of US-08-249-182-2 (1-6) with:
Data bank : A-GeneSeq 15, all entries



SCORE 0	1	1	2	2	3	3	4	4	5
STDEV 1	2	3	4						

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.10

Tines:	CPU	Total Elapsed
	00:00:27.91	00:00:30.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	3792

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
1. R37444	Autotaxin peptide ATX 19.	7	5	5	4.55	0
2. R52633	Guinea pig PH-30, 30 alpha su	289	5	5	4.55	0
3. R39354	EpiP protein.	461	5	5	4.55	0
4. R41043	CD4-EBA175 fusion protein.	1786	5	5	4.55	0
**** 3 standard deviations above mean ****						
5. R31940	In vivo tumour binding peptid	5	4	4	3.64	0
6. P50562	Sequence of peptide with immu	5	4	4	3.64	0
7. R24686	Immunomodulatory peptide.	6	4	4	3.64	0
8. R24619	Immunomodulatory peptide.	6	4	4	3.64	0
9. R24632	Immunomodulatory peptide.	6	4	4	3.64	0
10. R24639	Immunomodulatory peptide.	7	4	4	3.64	0
11. R24589	Immunomodulatory peptide.	7	4	4	3.64	0
12. R24588	Immunomodulatory peptide.	7	4	4	3.64	0
13. R21056	Gamma-carboxylase, N-terminus	10	4	4	3.64	0
14. R30889	Cell adhesion polypeptide.	19	4	4	3.64	0
15. R26829	Cell adhesion polypeptide.	19	4	4	3.64	0
16. R13381	Vascular permeability factor	21	4	4	3.64	0
17. R36681	Guinea pig VPF N-terminal.	25	4	4	3.64	0
18. R11665	N-terminal sequence of vascul	36	4	4	3.64	0
19. P70481	Amino-terminal amino acid seq	52	4	4	3.64	0
20. P70669	Sequence of N-terminal of bov	56	4	4	3.64	0
21. P82904	Acetylcholinesterase-like pro	81	4	4	3.64	0

22. R03700	Hsp-37 protein.	99	4	4	3.64	0
23. R22364	GroES structural protein.	102	4	4	3.64	0
24. R29702	Tapetum protein from A9 gene	105	4	4	3.64	0
25. P81001	Sequence (I) of human granulo	128	4	4	3.64	0
26. R34496	Bovine acidic Fibroblast Grow	140	4	4	3.64	0
27. R25915	Human acidic fibroblast growt	140	4	4	3.64	0
28. R13030	Brain-derived acidic fibrobla	140	4	4	3.64	0
29. P90069	Bovine acidic fibroblast grow	140	4	4	3.64	0
30. R25569	Recombinant bovine Ala47, Gly	141	4	4	3.64	0
31. R34010	Bovine lysozyme c.	145	4	4	3.64	0
32. P92066	Amino acid sequence of bovine	145	4	4	3.64	0
33. P91300	Signal sequence fused to the	147	4	4	3.64	0
34. R14626	Beta-lactoglobulin contg. pos	162	4	4	3.64	0
35. R42169	Haemopoietic stem cell promot	188	4	4	3.64	0
36. R41349	NXG2 encoded xyloglucanase.	190	4	4	3.64	0
37. P91467	Peptide with firin-binding ac	192	4	4	3.64	0
38. R26064	Human FcER1 alpha-subunit and	235	4	4	3.64	0
39. R47462	Truncated xylanase (XYLA).	236	4	4	3.64	0
40. R47457	Truncated xylanase (XYLA).	241	4	4	3.64	0

1. US-08-249-182-2 (1-6)

R37444 Autotaxin peptide ATX 19.

ID R37444 standard; peptide; 7 AA.
AC R37444;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 19.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
FH Key Location/Qualifiers
FT Modified_site 2
FT /note= "potentially glycosylated residue"
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 19. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 7 AA;
SQ 0 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 0 T; 0 W; 1 Y; 1 V;
SQ 1 Others;

Initial Score = 5 Optimized Score = 5 Significance = 4.55
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
|||||
PXLDVYK

2. US-08-249-182-2 (1-6)

R52633 Guinea pig PH-30, 30 alpha subunit.

ID R52633 standard; Protein; 289 AA.
 AC R52633;
 DT 22-JUN-1994 (first entry)
 DE Guinea pig PH-30, 30 alpha subunit.
 KW PH-20; PH-30; contraceptive; fertilisation; sperm surface protein;
 KW vaccine; sperm-egg fusion.
 OS Cavia cobaya.
 PN W09325233-A.
 PD 23-DEC-1993.
 PF 10-JUN-1993; U05640.
 PR 12-JUN-1992; US-897883.
 PA (UYCO-) UNIV CONNECTICUT.
 PI Myles DG, Primakoff P;
 DR WPI; 94-007200/01.
 DR N-PSDB; Q54636.
 PT Contraceptive vaccine for reducing sperm-egg fusion - comprises
 PT peptide from sperm surface protein which stimulates antibody
 PT prodn.
 PS Example 5; Fig 8B; 79pp; English.
 CC Sperm surface proteins or peptides stimulate an immune response to
 CC produce antibodies which block sperm-egg fusion and provide
 CC contraception. Pref. sperm surface proteins are the PH-20 and PH-30
 CC sperm surface proteins.
 SQ Sequence 289 AA;
 SQ 26 A; 2 R; 14 N; 12 D; 0 B; 23 C; 8 Q; 26 E; 0 Z; 32 G; 4 H;
 SQ 19 I; 17 L; 9 K; 1 M; 8 F; 29 P; 24 S; 22 T; 2 W; 5 Y; 6 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.55
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X   X
      PLDVYK
      ||| ||
YCTGQSGKCP LDTYKQDGT PCNEGFFCVSKGCTDPGIQCATYFGHGARSAPDACYTTLNSIGNIF
      10   X   20       30       40       50       60
  
```

3. US-08-249-182-2 (1-6)

R39354 EpiP protein.

ID R39354 standard; Protein; 461 AA.
 AC R39354;
 DT 15-SEP-1993 (first entry)
 DE EpiP protein.
 KW Epidermin; derivatives; lantibiotic.
 OS Staphylococcus epidermis.
 PN EP-543195-A.
 PD 26-MAY-1993.
 PF 30-OCT-1992; 118598.
 PR 31-OCT-1991; US-784234.
 PA (THOM) THOMAE GMBH KARL.
 PI Augustin J, Engelke G, Entian K, Gotz F, Jung G, Kaletta C, Klein C;
 PI Kellner R, Kupke T, Rosenstein R, Schnell N, Wieland B.
 DR WPI; 93-168917/21.
 DR N-PSDB; Q42541.
 PT Novel DNA molecule - encoding Epi B, C, D, P or Q enzymes
 PT involved in biosynthesis of lantibiotic epidermin.
 PS Claim 5; Fig 9; 52pp; English.
 CC The sequence is that of EpiP which is believed to be responsible

CC for cleaving mature epidermin from the N-terminal leader peptide,
 CC based on its striking homologies with the essential domain of
 CC serine proteases.
 SQ Sequence 461 AA;
 SQ 21 A; 8 R; 47 N; 30 D; 0 B; 2 C; 13 Q; 26 E; 0 Z; 32 G; 5 H;
 SQ 27 I; 38 L; 67 K; 8 M; 15 F; 8 P; 40 S; 16 T; 3 W; 22 Y; 33 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.55
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||||
 SLAAPKVSGALALEIDKYQLKDQPETAIELFKKKGIEKEKYMDKKHYGNGLDVYKLLKE
 410 420 430 440 450 X 460

4. US-08-249-182-2 (1-6)

R41043 CD4-EBA175 fusion protein.

ID R41043 standard; protein; 1786 AA.
 AC R41043;
 DT 22-MAR-1994 (first entry)
 DE CD4-EBA175 fusion protein.
 KW Merozoite; Erythrocyte Binding Antigen 175; malaria; HIV; env;
 KW human immunodeficiency virus; envelope glycoprotein; hybrid protein;
 KW red blood cell; erythrocyte; AIDS; molecular machine.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 FH Key Location/Qualifiers
 FT Region 1..371
 FT /note= "residues 1-371 of CD4"
 FT Region 372..1786
 FT /note= "residues 20-1435 of EBA-175"
 PN W09318160-A.
 PD 16-SEP-1993.
 PF 10-MAR-1993; G00505.
 PR 11-MAR-1992; GB-005276.
 PR 08-JUL-1992; GB-014481.
 PR 24-JUL-1992; GB-015829.
 PR 16-SEP-1992; GB-019562.
 PR 03-MAR-1993; GB-004311.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;
 DR WPI; 93-303474/38.
 PT Anti-viral fusion peptide(s) - comprise viral-binding component
 PT and malaria merozoite red cell binding component, for treating
 PT e.g. HIV, and hepatitis
 PS Claim 9; Page 44-47; 69pp; English.
 CC The hybrid protein NH2-CD4(1-371)-EBA175(20-1435)-COOH is a
 CC specifically claimed example of a fusion protein of the invention;
 CC it comprises at least part of the CD4 molecule fused to a peptide
 CC from a malarial parasite merozoite protein with affinity for red
 CC blood cells. The fusion protein can bind free HIV in the blood to
 CC red blood cells and consequently reduce viral titre, prevent
 CC transmission of the virus and improve safety of blood transfusions.
 SQ Sequence 1786 AA;
 SQ 47 A; 65 R; 163N; 126D; 0 B; 41 C; 63 Q; 169E; 0 Z; 73 G; 40 H;
 SQ 103I; 134L; 202K; 29 M; 55 F; 51 P; 160S; 98 T; 28 W; 54 Y; 85 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.55
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

PLDVYK
 VVVLNPEAGMWGCLLSDSGQVLLESNIKVLPTWSTPVKARNEYDIKENEKFLDVYKEKFNELDKKKYGNVQK
 340 350 360 370 380 X 390 400

TDKKIFTFIENKLDILNNSKFNKRWKSYPGTPDNI
 410 420 430 440

5. US-08-249-182-2 (1-6)

R31940 In vivo tumour binding peptide contg. Leu-Asp-Val.

ID R31940 standard; peptide; 5 AA.
 AC R31940;
 DT 04-JUN-1993 (first entry)
 DE In vivo tumour binding peptide contg. Leu-Asp-Val.
 KW Diagnosis; in vivo tumour imaging; therapy; treatment; cancer; LDV.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 2..4
 FT /note= "LDV - binds to tumour LDV binding sites"
 PN EP-527056-A.
 PD 10-FEB-1993.
 PF 06-AUG-1992; 307198.
 PR 06-AUG-1991; GB-016925.
 PA (ANTI-) ANTISOMA LTD.
 PI Stuttle AWJ.
 DR WPI; 93-046975/06.
 PT Peptide(s) contg. leucine-aspartic acid-valine sequence, bind
 PT tumours in vivo - for diagnosing and treating tumours
 PS Claim 3; Page 9; 13pp; English.
 CC The peptide contains the sequence Leu-Asp-Val (LDV) and as a result
 CC is capable of binding in vivo with pathological tissues such as
 CC tumours contg. an LDV binding site. It may have an attached or
 CC conjugated radioactive label or cytotoxin. It can be used in in vivo
 CC tumour imaging, diagnosis and treatment of tumours. Using the peptide
 CC the disadvantages of monoclonal antibody based cytotoxic and
 CC diagnostic reagents are overcome. It has specific binding properties
 CC to tumour associated binding sites and is rapidly transported to the
 CC tumour site following injection. It has rapid clearance of unbound
 CC reagent from the body following admin. thereby considerably reducing
 CC treatment times and reducing effective dosage of potentially highly
 CC toxic reagents whether administered for diagnostic or therapeutic
 CC purposes.
 SQ Sequence 5 AA;
 SQ 0 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 2 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 YLDVY
 X X

6. US-08-249-182-2 (1-6)

P50562 Sequence of peptide with immunomodulatory activity

ID P50562 standard; peptide; 5 AA.
 AC P50562;
 DT 29-NOV-1991 (first entry)
 DE Sequence of peptide with immunomodulatory activity.

KW immunomodulatory; lymphopoietic stem cell differentiation;
 KW haemopoiesis; DiGeorge syndrome; therapy.
 FH Key Location/Qualifiers
 FT Modified -site 1
 FT /label= N-alpha-acetyl-Arg
 FT Modified -site 2
 FT /label= 4-methyl-Leu
 FT Modified -site 5
 FT /label= Tyr-OH
 PN US4505853-A.
 PD 19-MAR-1985.
 PF 18-NOV-1983; 553281.
 PR 18-NOV-1983; US-553281.
 PA (ORTH) ORTHO PHARM CORP.
 PI Goldstein G, Heavner G, Kroon D, Audhya T;
 DR WPI; 85-086695/14.
 PT New peptide(s) - useful immuno-regulatory agents with improved
 PT resistance to enzymatic degradation in body
 PS Claim 28; column 38; 20pp; English.
 CC The peptides of the invention induce differentiation of
 CC lymphopoietic stem cells from haemopoietic tissues into thymus-
 CC derived cells and so are useful for treating such conditions as
 CC DiGeorge syndrome. They also assist collective body immunity and so
 CC can be used with autoimmune diseases, etc. Dose is 10-100
 CC micrograms/kg daily for DiGeorge syndrome.
 SQ Sequence 5 AA;
 SQ 0 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 RLDVY
 X X

7. US-08-249-182-2 (1-6)

R24686 Immunomodulatory peptide.

ID R24686 standard; peptide; 6 AA.
 AC R24686;
 DT 03-DEC-1992 (first entry)
 DE Immunomodulatory peptide.
 KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
 KW inflammation; wounds; lymphocyte; vaccine.
 OS Synthetic.
 PN W09209628-A.
 PD 11-JUN-1992.
 PF 22-NOV-1991; U08795.
 PR 23-NOV-1990; US-617494.
 PA (IMMU-) IMMUNODYNAMICS INC.
 PI Atkin A;
 DR WPI; 92-217021/26.
 PT New synthetic immunomodulatory peptide(s) - for treating
 PT immunodeficiencies, immunosuppression and T-cell subset
 PT deviations and immuno-therapy of infections, inflammation, wounds
 PT etc.
 PS Claim 10; Page 36; 52pp; English.
 CC The immunomodulatory peptide is a specific example of a peptide cpd.
 CC (or an acid or base salt) constructed by combination and/or
 CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
 CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,

CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
 CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
 CC be used for immunomodulation of various immunodeficiencies and
 CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
 CC enhancement of a vaccines efficacy, as well as for immunotherapy,
 CC including infections, local or systemic complications of non-
 CC infectious diseases, postoperative inflammations, wounds and burns.
 CC See also R24583-R24701.

SQ Sequence 6 AA;
 SQ 0 A; 1 R; 1 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 1 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 RNDVYK
 X X

8. US-08-249-182-2 (1-6)
 R24619 Immunomodulatory peptide.

ID R24619 standard; peptide; 6 AA.
 AC R24619;
 DT 03-DEC-1992 (first entry)
 DE Immunomodulatory peptide.
 KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
 KW inflammation; wounds; lymphocyte; vaccine.
 OS Synthetic.
 PN W09209628-A.
 PD 11-JUN-1992.
 PF 22-NOV-1991; U08795.
 PR 23-NOV-1990; US-617494.
 PA (IMMU-) IMMUNODYNAMICS INC.
 PI Atkin A;
 DR WPI; 92-217021/26.
 PT New synthetic immunomodulatory peptide(s) - for treating
 PT immunodeficiencies, immunosuppression and T-cell subset
 PT deviations and immuno-therapy of infections, inflammation, wounds
 PT etc.
 PS Claim 10; Page 36; 52pp; English.
 CC The immunomodulatory peptide is a specific example of a peptide cpd.
 CC (or an acid or base salt) constructed by combination and/or
 CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
 CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,
 CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
 CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
 CC be used for immunomodulation of various immunodeficiencies and
 CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
 CC enhancement of a vaccines efficacy, as well as for immunotherapy,
 CC including infections, local or systemic complications of non-
 CC infectious diseases, postoperative inflammations, wounds and burns.
 CC See also R24583-R24701.

SQ Sequence 6 AA;
 SQ 0 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 2 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X

PLDVYK
||||
RKDVYK
X X

9. US-08-249-182-2 (1-6)

R24632 Immunomodulatory peptide.

ID R24632 standard; peptide; 6 AA.
AC R24632;
DT 03-DEC-1992 (first entry)
DE Immunomodulatory peptide.
KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
KW inflammation; wounds; lymphocyte; vaccine.
OS Synthetic.
PN W09209628-A.
PD 11-JUN-1992.
PF 22-NOV-1991; U08795.
PR 23-NOV-1990; US-617494.
PA (IMMU-) IMMUNODYNAMICS INC.
PI Atkin A;
DR WPI; 92-217021/26.
PT New synthetic immunomodulatory peptide(s) - for treating
PT immunodeficiencies, immunosuppression and T-cell subset
PT deviations and immuno-therapy of infections, inflammation, wounds
PT etc.
PS Claim 10; Page 36; 52pp; English.
CC The immunomodulatory peptide is a specific example of a peptide cpd.
CC (or an acid or base salt) constructed by combination and/or
CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,
CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
CC be used for immunomodulation of various immunodeficiencies and
CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
CC enhancement of a vaccines efficacy, as well as for immunotherapy,
CC including infections, local or systemic complications of non-
CC infectious diseases, postoperative inflammations, wounds and burns.
CC See also R24583-R24701.
SQ Sequence 6 AA;
SQ 0 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 2 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 2 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||||
YKDVYK
X X

10. US-08-249-182-2 (1-6)

R24639 Immunomodulatory peptide.

ID R24639 standard; peptide; 7 AA.
AC R24639;
DT 03-DEC-1992 (first entry)
DE Immunomodulatory peptide.
KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
KW inflammation; wounds; lymphocyte; vaccine.
OS Synthetic.
PN W09209628-A.

PD 11-JUN-1992.
 PF 22-NOV-1991; U08795.
 PR 23-NOV-1990; US-617494.
 PA (IMMU-) IMMUNODYNAMICS INC.
 PI Atkin A;
 DR WPI; 92-217021/26.
 PT New synthetic immunomodulatory peptide(s) - for treating
 PT immunodeficiencies, immunosuppression and T-cell subset
 PT deviations and immuno-therapy of infections, inflammation, wounds
 PT etc.
 PS Claim 10; Page 36; 52pp; English.
 CC The immunomodulatory peptide is a specific example of a peptide cpd.
 CC (or an acid or base salt) constructed by combination and/or
 CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
 CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,
 CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
 CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
 CC be used for immunomodulation of various immunodeficiencies and
 CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
 CC enhancement of a vaccines efficacy, as well as for immunotherapy,
 CC including infections, local or systemic complications of non-
 CC infectious diseases, postoperative inflammations, wounds and burns.
 CC See also R24583-R24701.
 SQ Sequence 7 AA;
 SQ 0 A; 1 R; 1 N; 1 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 1 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 RDVYKQN
 X X

11. US-08-249-182-2 (1-6)
 R24589 Immunomodulatory peptide.

ID R24589 standard; peptide; 7 AA.
 AC R24589;
 DT 03-DEC-1992 (first entry)
 DE Immunomodulatory peptide.
 KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
 KW inflammation; wounds; lymphocyte; vaccine.
 OS Synthetic.
 PN W09209628-A.
 PD 11-JUN-1992.
 PF 22-NOV-1991; U08795.
 PR 23-NOV-1990; US-617494.
 PA (IMMU-) IMMUNODYNAMICS INC.
 PI Atkin A;
 DR WPI; 92-217021/26.
 PT New synthetic immunomodulatory peptide(s) - for treating
 PT immunodeficiencies, immunosuppression and T-cell subset
 PT deviations and immuno-therapy of infections, inflammation, wounds
 PT etc.
 PS Claim 9; Page 34; 52pp; English.
 CC The immunomodulatory peptide is a specific example of a peptide cpd.
 CC (or an acid or base salt) constructed by combination and/or
 CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
 CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,
 CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
 CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may

CC be used for immunomodulation of various immunodeficiencies and
CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
CC enhancement of a vaccines efficacy, as well as for immunotherapy,
CC including infections, local or systemic complications of non-
CC infectious diseases, postoperative inflammations, wounds and burns.
CC See also R24583-R24701.

SQ Sequence 7 AA;
SQ 0 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 3 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||||
KEKDVYK
X X

12. US-08-249-182-2 (1-6)

R24588 Immunomodulatory peptide.

ID R24588 standard; peptide; 7 AA.
AC R24588;
DT 03-DEC-1992 (first entry)
DE Immunomodulatory peptide.
KW Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
KW inflammation; wounds; lymphocyte; vaccine.
OS Synthetic.
PN W09209628-A.
PD 11-JUN-1992.
PF 22-NOV-1991; U08795.
PR 23-NOV-1990; US-617494.
PA (IMMU-) IMMUNODYNAMICS INC.
PI Atkin A;
DR WPI; 92-217021/26.
PT New synthetic immunomodulatory peptide(s) - for treating
PT immunodeficiencies, immunosuppression and T-cell subset
PT deviations and immuno-therapy of infections, inflammation, wounds
PT etc.
PS Claim 9; Page 34; 52pp; English.
CC The immunomodulatory peptide is a specific example of a peptide cpd.
CC (or an acid or base salt) constructed by combination and/or
CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,
CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;
CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
CC be used for immunomodulation of various immunodeficiencies and
CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
CC enhancement of a vaccines efficacy, as well as for immunotherapy,
CC including infections, local or systemic complications of non-
CC infectious diseases, postoperative inflammations, wounds and burns.
CC See also R24583-R24701.

SQ Sequence 7 AA;
SQ 0 A; 0 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 3 K; 0 M; 0 F; 0 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
Residue Identity = 66% Matches = 4 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||||

13. US-08-249-182-2 (1-6)

R21056 Gamma-carboxylase, N-terminus.

ID R21056 standard; Protein; 10 AA.
 AC R21056;
 DT 01-JUN-1992 (first entry)
 DE Gamma-carboxylase, N-terminus.
 KW Degenerate; Vitamin K dependent proteins; PCR.
 OS Homo sapiens.
 PN W09201795-A.
 PD 06-FEB-1992.
 PF 22-JUL-1991; U05177.
 PR 23-JUL-1990; US-557220.
 PR 14-MAR-1991; US-669735.
 PA (ZYMO-) ZYMOGENETICS INC.
 PI Berkner KL;
 DR WPI; 92-064951/08.
 PT Gamma-carboxylase protein compsns. - used in recombinant prodn.
 PT of active vitamin=K dependent proteins
 PS Claim 6; Table 9; 91pp; English.
 CC Nucleotide sequences encoding gamma-carboxylase were obtd. using PCR,
 CC and oligonucleotides designed from amino acid sequences determined
 CC by microsequencing of partially purified material. SEQ ID No 20
 CC (peptide 6) is one of seven alternative sequences for the N-terminus
 CC of gamma carboxylase, due to the degeneracy of the genetic code.
 CC Obtaining the full DNA and protein sequence of gamma-carboxylase
 CC will allow proteins such as Factor VII, Factor IX, prothrombin,
 CC (activated) protein C, protein S, protein Z, or osteocalcin to be
 CC easily produced by recombinant techniques.
 CC See also R21049-55, R23010.
 SQ Sequence 10 AA;
 SQ 0 A; 0 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 0 I; 2 L; 1 K; 0 M; 0 F; 1 P; 1 S; 1 T; 0 W; 1 Y; 0 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 || ||
 TLPSGLDDYK
 X 10

14. US-08-249-182-2 (1-6)

R30889 Cell adhesion polypeptide.

ID R30889 standard; peptide; 19 AA.
 AC R30889;
 DT 09-FEB-1993 (first entry)
 DE Cell adhesion polypeptide.
 KW MOLT-4; human; lymphoblastic leukaemia; A375-SM; metastatic;
 KW melanoma; H1080; fibrosarcoma; LDV; LDL; IDA; inflammatory disease;
 KW rheumatoid arthritis; asthma; sepsis; graft rejection; reperfusion.
 OS Synthetic.
 PN W09213887-A.
 PD 20-AUG-1992.
 PF 06-FEB-1992; G00226.
 PR 07-FEB-1991; GB-002655.
 PR 08-FEB-1991; GB-002818.
 PA (UYMA-) UNIV VICTORIA MANCHESTER.

PI Humphries MJ;
 DR WPI; 92-299988/36.
 PT New cell adhesion (poly)peptide(s) modifying cell adhesive
 PT properties - useful in treating inflammatory conditions e.g.
 PT rheumatoid arthritis, asthma, inflammatory bowel disease, sepsis,
 PT etc.
 PS Disclosure; Page 4; 23pp; English.
 CC The peptide is an example of a cell adhesion polypeptide contg. the
 CC amino sequence X-Asp-Y-(A)n-Phe, where X and Y = Ala, Leu, Ile or
 CC Val, A= any amino acid and n= 3-10. At least a subsequence of the
 CC polypeptide is adherent for MOLT-4 human lymphoblastic leukaemia,
 CC A375-SM human metastatic melanoma or H1080 human fibrosarcoma cells.
 CC The cell adhesion peptides are used to modify or control the
 CC adhesive properties of cells, e.g. in treatment of inflammatory
 CC conditions such as rheumatoid arthritis, asthma, sepsis, graft
 CC rejection, inflammatory bowel disease, reperfusion of cardiac tissue
 CC after myocardial infarction, and coagulatory disorders. They are
 CC selective antagonists of cell adhesion, e.g. they promote adhesion
 CC of the specified cells but inhibit adhesion to the natural adhesion
 CC protein contg. the adhesive sequence.
 CC See also R26821-30 and R30887-903.
 SQ Sequence 19 AA;
 SQ 2 A; 0 R; 0 N; 3 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 2 G; 0 H;
 SQ 2 I; 1 L; 0 K; 0 M; 2 F; 3 P; 0 S; 0 T; 0 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 PIIDVAPLDVGAPDGEFGF
 X 10 X

15. US-08-249-182-2 (1-6)
 R26829 Cell adhesion polypeptide.

ID R26829 standard; peptide; 19 AA.
 AC R26829;
 DT 09-FEB-1993 (first entry)
 DE Cell adhesion polypeptide.
 KW MOLT-4; human; lymphoblastic leukaemia; A375-SM; metastatic;
 KW melanoma; H1080; fibrosarcoma; LDV; LDL; IDA; inflammatory disease;
 KW rheumatoid arthritis; asthma; sepsis; graft rejection; reperfusion.
 OS Synthetic.
 PN W09213887-A.
 PD 20-AUG-1992.
 PF 06-FEB-1992; G00226.
 PR 07-FEB-1991; GB-002655.
 PR 08-FEB-1991; GB-002818.
 PA (UYMA-) UNIV VICTORIA MANCHESTER.
 PI Humphries MJ;
 DR WPI; 92-299988/36.
 PT New cell adhesion (poly)peptide(s) modifying cell adhesive
 PT properties - useful in treating inflammatory conditions e.g.
 PT rheumatoid arthritis, asthma, inflammatory bowel disease, sepsis,
 PT etc.
 PS Disclosure; Page 4; 23pp; English.
 CC The peptide is an example of a cell adhesion polypeptide contg. the
 CC amino sequence X-Asp-Y-(A)n-Phe, where X and Y = Ala, Leu, Ile or
 CC Val, A= any amino acid and n= 3-10. At least a subsequence of the
 CC polypeptide is adherent for MOLT-4 human lymphoblastic leukaemia,
 CC A375-SM human metastatic melanoma or H1080 human fibrosarcoma cells.
 CC The cell adhesion peptides are used to modify or control the

CC adhesive properties of cells, e.g. in treatment of inflammatory
 CC conditions such as rheumatoid arthritis, asthma, sepsis, graft
 CC rejection, inflammatory bowel disease, reperfusion of cardiac tissue
 CC after myocardial infarction, and coagulatory disorders. They are
 CC selective antagonists of cell adhesion, e.g. they promote adhesion
 CC of the speccified cells but inhibit adhesion to the natural adhesion
 CC protein contg. the adhesive sequence.
 CC See also R26821-30 and R30887-903.

SQ Sequence 19 AA;
 SQ 1 A; 0 R; 0 N; 2 D; 0 B; 1 C; 1 Q; 0 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 4 L; 0 K; 0 M; 1 F; 2 P; 4 S; 0 T; 0 W; 0 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.64
 Residue Identity = 66% Matches = 4 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 CSQPLDVILLLDGSSSFPA
 X 10

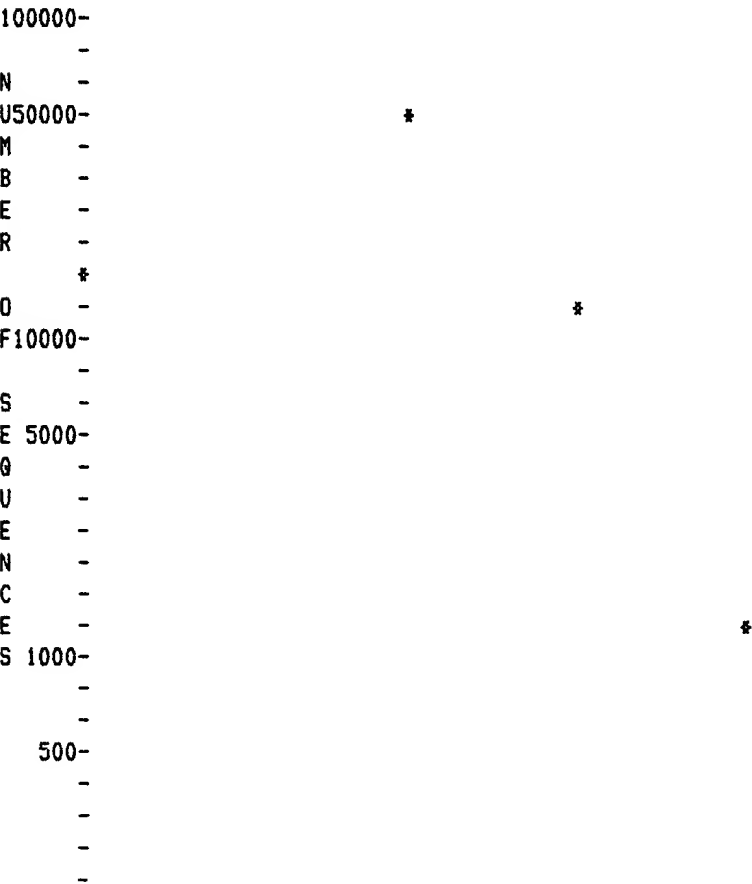
> 0 <
 0| |0 IntelliGenetics
 > 0 <

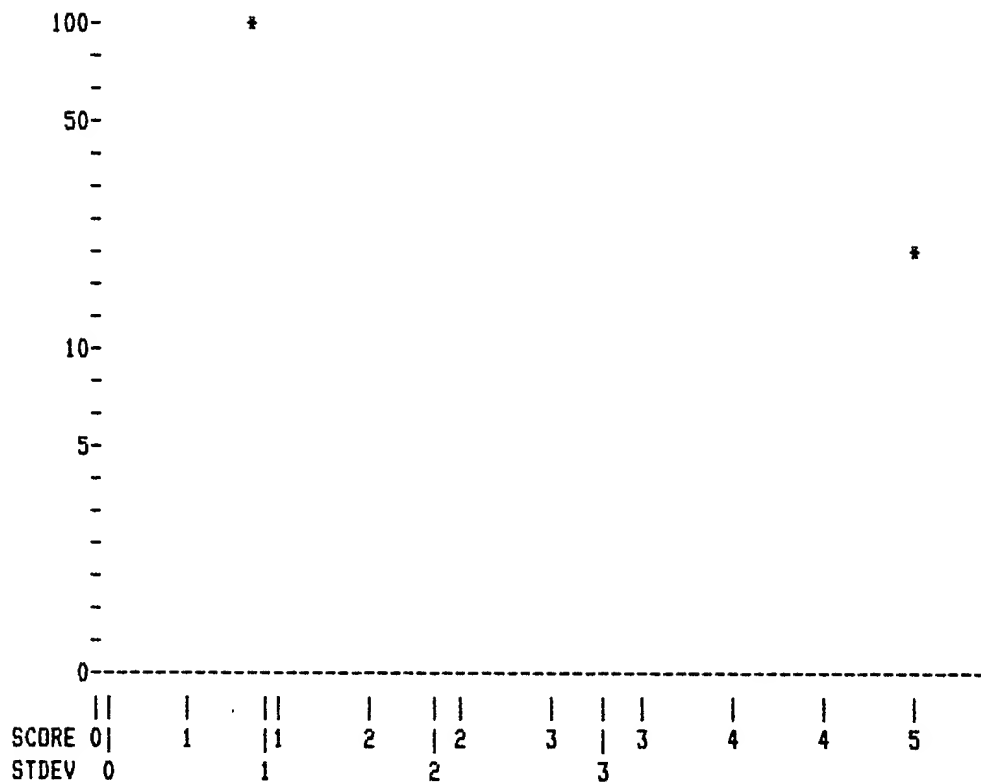
FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_2p.res made by on Thu 22 Sep 94 10:50:46-PDT.

Query sequence being compared:US-08-249-182-2 (1-6)
 Number of sequences searched: 70848
 Number of scores above cutoff: 4700

Results of the initial comparison of US-08-249-182-2 (1-6) with:
 Data bank : PIR 41, all entries





PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.98

Times:	CPU	Total Elapsed
	00:01:23.97	00:01:32.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	4700

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

**** 4 standard deviations above mean ****						
1. A42329	autotaxin - human (fragments)	114	5	5	4.09	0
2. S19501	hypothetical protein YCR086W	190	5	5	4.09	0
3. S23402	sperm surface protein PH-30 a	289	5	5	4.09	0
4. A24849	ADP,ATP carrier protein AAC1	309	5	5	4.09	0
5. JT0584	deoxyribonuclease I (EC 3.1.2	327	5	5	4.09	0
6. S23420	hypothetical protein - Staphy	461	5	5	4.09	0
7. S16250	phytoene desaturase - Synecho	474	5	5	4.09	0
8. S35519	hypothetical protein - Caenor	537	5	5	4.09	0
9. UFECAG	fumarate hydratase (EC 4.2.1.	548	5	5	4.09	0
10. JT0742	tartronate-semialdehyde synth	593	5	5	4.09	0
11. S03208	type III site-specific deoxyr	645	5	5	4.09	0
12. OYRTA1	guanylate cyclase (EC 4.6.1.2	690	5	5	4.09	0
13. OYBD77	guanylate cyclase (EC 4.6.1.2	691	5	5	4.09	0
14. S23098	guanylate cyclase (EC 4.6.1.2	717	5	5	4.09	0
15. A42163	Na+/myo-inositol cotransporte	718	5	5	4.09	0
16. S23019	DNA-directed DNA polymerase (882	5	5	4.09	0
17. S11561	EBA-175 protein - Plasmodium	1426	5	5	4.09	0
18. A37793	erythrocyte-binding antigen 1	1435	5	5	4.09	0
19. S34670	splicing factor PRP8 - yeast	2413	5	5	4.09	0
**** 3 standard deviations above mean ****						
20. A34477	heart-derived growth factor -	26	4	4	3.07	0
21. A42272	brain-type creatine kinase, p	28	4	4	3.07	0
22. C25532	exonuclease V alpha chain - E	31	4	4	3.07	0
23. A60706	vascular endothelial growth f	36	4	4	3.07	0
24. SDDVEG	desulforedoxin - Desulfovibri	37	4	4	3.07	0
25. D45731	comC-alpha 3'-region hypothet	45	4	4	3.07	0
26. S16198	diuretic peptide - house cric	46	4	4	3.07	0
27. A48542	CRF-related diuretic peptide	46	4	4	3.07	0
28. A29180	cytochrome-c oxidase (EC 1.9.	46	4	4	3.07	0
29. S09354	testis-determining factor ZFY	49	4	4	3.07	0
30. A35722	creatine kinase (EC 2.7.3.2)	52	4	4	3.07	0
31. R3KM72	ribosomal protein S7-2 - Chla	52	4	4	3.07	0
32. E38269	protein-tyrosine kinase (EC 2	57	4	4	3.07	0
33. A22810	small acid-soluble spore prot	65	4	4	3.07	0
34. S29406	photosystem II 10K protein -	68	4	4	3.07	0
35. A35537	cytochrome-c oxidase (EC 1.9.	70	4	4	3.07	0
36. A38646	osteopontin-related 20K prote	75	4	4	3.07	0
37. B35540	cruciferin alpha-2 chain/beta	76	4	4	3.07	0
38. JE0003	hypothetical 8.6K protein - p	76	4	4	3.07	0
39. A03082	villin - chicken (fragment)	76	4	4	3.07	0
40. A25703	villin - chicken (fragment)	76	4	4	3.07	0

1. US-08-249-182-2 (1-6)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329
 ##status preliminary
 ##molecule_type protein
 ##residues 1-114 ##label STR
 ##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;

##note sequence extracted from NCB1 backbone
 SUMMARY #length 114 #checksum 7335
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||||

GQPLWITATKSPFFENINLYYDVPWNETIPEEVTXPNYLQAEVSYPAFKPLDVYKWHVAAN
 60 70 80 90 100 X 110

2. US-08-249-182-2 (1-6)

S19501 hypothetical protein YCR086W - yeast (Saccharomyce

ENTRY S19501 #type complete
 TITLE hypothetical protein YCR086W - yeast (Saccharomyces cerevisiae)
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 18-Jun-1993
 ACCESSIONS S19501
 REFERENCE S19351
 #authors Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippsen, P.; Schweitzer, B.; Spiegelberg, R.; Steiner, S.
 #submission submitted to the Protein Sequence Database, March 1992
 #accession S19501
 ##molecule_type DNA
 ##residues 1-190 ##label DUS
 ##cross-references EMBL:X59720

GENETICS

#map_position 3R

SUMMARY #length 190 #molecular-weight 21751 #checksum 6551
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||||

MDPLTVYKNSVKQGDIDSDLVLANLVNENFVLSEKLDTKATEIKQLQKQIDSLNAQVK
 X 10 20 30 40 50

3. US-08-249-182-2 (1-6)

S23402 sperm surface protein PH-30 alpha chain - guinea p

ENTRY S23402 #type complete
 TITLE sperm surface protein PH-30 alpha chain - guinea pig
 ORGANISM #formal_name Cavia porcellus #common_name guinea pig
 DATE 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 18-Jun-1993
 ACCESSIONS S23402; S25695
 REFERENCE S23402
 #authors Blobel, C.P.; Wolfsberg, T.G.; Turck, C.W.; Myles, D.G.; Prinakoff, P.; White, J.M.
 #journal Nature (1992) 356:248-252
 #title A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion.

```

#accession      S23402
##molecule_type mRNA
##residues      1-289 ##label BLD
##cross-references EMBL:Z11719
#accession      S25695
##molecule_type protein
##residues      1,'X',3-8,'X',10-18;103-115 ##label BLD1
KEYWORDS        glycoprotein; membrane protein
FEATURE
  228-252        #domain transmembrane #status predicted #label TMM\
  73,161         #binding_site carbohydrate (Asn) (covalent) #status
                  predicted
SUMMARY         #length 289 #molecular-weight 29735 #checksum 7298
SEQUENCE

```

```

Initial Score   =      5  Optimized Score =      5  Significance =  4.09
Residue Identity =    83%  Matches         =      5  Mismatches  =      1
Gaps            =      0  Conservative Substitutions =      0

```

```

      X   X
      PLDVYK
      ||| ||
YCTGQSGKCPLDTYKQDGTPCNEGFFCVSKGCTDPGIQCATYFGHGARSAPDACYTTLNSIGNIF
      10   X   20       30       40       50       60

```

4. US-08-249-182-2 (1-6)

A24849 ADP,ATP carrier protein AAC1 - yeast (*Saccharomyces*

```

ENTRY          A24849      #type complete
TITLE          ADP,ATP carrier protein AAC1 - yeast (Saccharomyces
                  cerevisiae)
ALTERNATE_NAMES ADP,ATP carrier protein PET9; ADP,ATP translocase; ADP,ATP
                  translocator protein
ORGANISM       #formal_name Saccharomyces cerevisiae
DATE          31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change
                  08-Apr-1994
ACCESSIONS     A24849; C39654; S19035
REFERENCE      A24849
  #authors     Adrian, G.S.; McCammon, M.T.; Montgomery, D.L.; Douglas, M.G.
  #journal     Mol. Cell. Biol. (1986) 6:626-634
  #title       Sequences required for delivery and localization of the
                  ADP/ATP translocator to the mitochondrial inner membrane.
  #cross-references MUID:87064348
  #accession   A24849
    ##molecule_type DNA
    ##residues  1-309 ##label ADR
    ##cross-references EMBL:M12514
REFERENCE      A39654
  #authors     Hoyt, M.A.; Totis, L.; Roberts, B.T.
  #journal     Cell (1991) 66:507-517
  #title       Saccharomyces cerevisiae genes required for cell cycle arrest
                  in response to loss of microtubule function.
  #cross-references MUID:91330299
  #accession   C39654
    ##molecule_type DNA
    ##residues  265-309 ##label HOY
    ##cross-references GB:M64706
GENETICS
  #gene        AAC1
  #map_position 2L
CLASSIFICATION #superfamily ADP,ATP carrier protein
KEYWORDS       membrane protein; mitochondrion
SUMMARY        #length 309 #molecular-weight 34120 #checksum 3531
SEQUENCE

```

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                ||||
YAKWFAGNLFSGGAAGGLSLLFVYSLDYARTRLAADARGSKSTSQRQFNGLLDVYKKTCLKTDGLLGLYRGFV
  120      130      140      150      160 X   170      180

PSVLGIIVYRGLYFGLYDSFKPVLLTGALEGSFV
  190      200      210

```

5. US-08-249-182-2 (1-6)

JT0584 deoxyribonuclease I (EC 3.1.21.1) precursor - Stre

ENTRY JT0584 #type complete
 TITLE deoxyribonuclease I (EC 3.1.21.1) precursor - Streptococcus
 "equisimilis"
 ALTERNATE_NAMES streptodornase
 ORGANISM #formal_name Streptococcus "equisimilis"
 DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
 28-Apr-1993
 ACCESSIONS JT0584
 REFERENCE JT0584
 #authors Wolinowska, R.; Ceglowski, P.; Kok, J.; Venema, G.
 #journal Gene (1991) 106:115-119
 #title Isolation, sequence and expression in Escherichia coli,
 Bacillus subtilis and Lactococcus lactis of the DNase
 (streptodornase)-encoding gene from Streptococcus
 equisimilis H46A.
 #cross-references MUID:92039051
 #contents Strain H46A
 #accession JT0584
 ##molecule_type DNA
 ##residues 1-327 ##label WOL
 ##cross-references ENBL:X17241
 GENETICS
 #gene sdc
 KEYWORDS hydrolase
 FEATURE
 1-35 #domain signal sequence #label SIG\
 36-308 #protein deoxyribonuclease I #status predicted #label
 DEO
 SUMMARY #length 327 #molecular-weight 36844 #checksum 1320
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                ||| |
YGEYKDYYTVIGESNIDQSAFPKIYKTTERRVYKGQGTSEKRVTVSDVVYNPLDGYKRSTGAYGVVTKDMIDM
  50      60      70      80      90 X   100      110

SKGYREKWETNPEPSGWFYFNRADNEEISEKEY
  120      130      140

```

6. US-08-249-182-2 (1-6)

S23420 hypothetical protein - Staphylococcus epidermidis

ENTRY S23420 #type complete
 TITLE hypothetical protein - Staphylococcus epidermidis
 ORGANISM #formal_name Staphylococcus epidermidis
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993
 ACCESSIONS S23420
 REFERENCE S23413
 #authors Schnell, N.; Engelke, G.; Augustin, J.; Rosenstein, R.; Ungermann, V.; Goetz, F.; Entian, K.D.
 #journal Eur. J. Biochem. (1992) 204:57-68
 #title Analysis of genes involved in the biosynthesis of lantibiotic epidermin.
 #cross-references MUID:92155237
 #accession S23420
 ##status preliminary
 ##residues 1-461 ##label SCH
 ##cross-references EMBL:X62386
 SUMMARY #length 461 #molecular-weight 51814 #checksum 7694
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                PLDVYK
                                |||||
SLAAPKVSGALALEIDKYQLKDQETAIELFKKKGIEKEKYMDKKHYGNKLDVYKLLKE
  410      420      430      440      450 X      460

```

7. US-08-249-182-2 (1-6)

S16250 phytoene desaturase - Synechococcus sp.

ENTRY S16250 #type complete
 TITLE phytoene desaturase - Synechococcus sp.
 ORGANISM #formal_name Synechococcus sp.
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change 21-Nov-1993
 ACCESSIONS S16250
 REFERENCE S16250
 #authors Chamovitz, D.; Pecker, I.; Hirschberg, J.
 #journal Plant Mol. Biol. (1991) 16:967-974
 #title The molecular basis of resistance to the herbicide norflurazon.
 #cross-references MUID:91322511
 #accession S16250
 ##status preliminary
 ##residues 1-474 ##label CHA
 ##cross-references EMBL:X55289
 SUMMARY #length 474 #molecular-weight 53295 #checksum 4126
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                PLDVYK
                                |||||
ELVFAPAKDWIGRSDIEDILAATMAEIEKLFPQHFGSENPRLRKYIVKTPLSVYKATPGRQGYRPGQASPI
  350      360      370      380      390      400 X      410      420

ANFFLTGDYTMQRYLASMEGAVLSCKLTAQAIIA
  430      440      450

```

8. US-08-249-182-2 (1-6)

S35519 hypothetical protein - *Caenorhabditis elegans*

ENTRY S35519 #type complete
 TITLE hypothetical protein - *Caenorhabditis elegans*
 ORGANISM #formal_name *Caenorhabditis elegans*
 DATE 09-Dec-1993; #sequence_revision 09-Dec-1993; #text_change 09-Dec-1993
 ACCESSIONS S35519
 REFERENCE S35519
 #authors Li, W.; Shaw, J.E.
 #journal Nucleic Acids Res. (1993) 21:59-67
 #title A variant Tc4 transposable element in the nematode *C.elegans* could encode a novel protein.
 #accession S35519
 ##status preliminary
 ##residues 1-537 ##label LIA
 ##cross-references EMBL:L00665
 SUMMARY #length 537 #molecular-weight 62032 #checksum 5804
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 VFIPSPKKLYIMLDSWPAFKDHTTIKNLVPNGHDVVIRNIPEHTTGMIGPLDVYWNAPWKSIIKKFTAYAL
 360 370 380 390 400 410 X 420 430
 RTQTDYVIAQRNNAICMVSPLYHQISAHFRPFL
 440 450 460

9. US-08-249-182-2 (1-6)

UFECA0 fumarate hydratase (EC 4.2.1.2), iron-dependent -

ENTRY UFECA0 #type complete
 TITLE fumarate hydratase (EC 4.2.1.2), iron-dependent - *Escherichia coli*
 ALTERNATE_NAMES fumarase
 ORGANISM #formal_name *Escherichia coli*
 DATE 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 03-Feb-1994
 ACCESSIONS A03531; PX0048
 REFERENCE A93515
 #authors Miles, J.S.; Guest, J.R.
 #journal Nucleic Acids Res. (1984) 12:3631-3642
 #title Complete nucleotide sequence of the fumarase gene *funA*, of *Escherichia coli*.
 #cross-references MUID:84221385
 #accession A03531
 ##molecule_type DNA
 ##residues 1-548 ##label MIL
 REFERENCE A92783
 #authors Guest, J.R.; Miles, J.S.; Roberts, R.E.; Woods, S.A.
 #journal J. Gen. Microbiol. (1985) 131:2971-2984
 #title The fumarase genes of *Escherichia coli*: location of the *funB* gene and discovery of a new gene (*funC*).
 #cross-references MUID:86142617
 #contents annotation; identification of the structural gene
 REFERENCE PX0048
 #authors Ueda, Y.; Yumoto, N.; Tokushige, M.; Fukui, K.; Ohya-Nishiguchi, H.

#journal J. Biol. Chem. (1993) 268:3911-3919
 #title Purification and characterization of two types of fumarase
 from *Escherichia coli*.
 #cross-references MUID:92011457
 #contents strain W
 #accession PX0048
 ##molecule_type protein
 ##residues 2-21 ##label UED
 COMMENT In *E. coli*, three fumarate hydratase genes (*fumA*, *fumB*, and *fumC*)
 have been reported. This protein, the *fumA*-encoded fumarate
 hydratase, is an Fe-dependent 4Fe-4S hydratase.
 GENETICS
 #gene *fumA*
 #map_position 35.5 min
 CLASSIFICATION #superfamily iron-dependent fumarate hydratase
 KEYWORDS 4Fe-4S; carbon-oxygen lyase; homodimer; hydro-lyase;
 iron-sulfur protein
 FEATURE
 2-548 #protein fumarate hydratase, iron-dependent #label MAT
 SUMMARY #length 548 #molecular-weight 60298 #checksum 4355
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||| ||
 VLPTCQDTGTAIIVGKKGQRVWTGGGDEAALARGVYNTYIEDNLRYSQNAPLDMYKEVNTGTNLPAQIDLYA
 110 120 130 140 150 X 160 170
 VDGDEYKFLCIAKGGGSANKTYLYQETKALLTPG
 180 190 200

10. US-08-249-182-2 (1-6)

JT0742 tartronate-semialdehyde synthase (EC 4.1.1.47) - E

ENTRY JT0742 #type complete
 TITLE tartronate-semialdehyde synthase (EC 4.1.1.47) - *Escherichia coli*
 ALTERNATE_NAMES glyoxylate carbo-lyase
 ORGANISM #formal_name *Escherichia coli*
 DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
 ACCESSIONS JT0742; PN0053
 REFERENCE JT0742
 #authors Chang, Y.Y.; Wang, A.Y.; Cronan Jr., J.E.
 #journal J. Biol. Chem. (1993) 268:3911-3919
 #title Molecular cloning, DNA sequencing, and biochemical analyses
 of *Escherichia coli* glyoxylate carboligase.
 #accession JT0742
 ##molecule_type DNA
 ##residues 1-593 ##label CHA
 ##cross-references GB:L03845
 #accession PN0053
 ##molecule_type protein
 ##residues 2-31 ##label CH2
 COMMENT This enzyme catalyzes the condensation of two molecules of
 glyoxylate to give tartronic semialdehyde.
 GENETICS
 #gene *gcl*
 KEYWORDS carbon-carbon lyase; carboxy-lyase
 FEATURE
 2-593 #protein tartronate-semialdehyde synthase #status

SUMMARY #length 593 #molecular-weight 64731 #checksum 8186
SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                || ||
AVTVREAALVPRVLQAFHLMRSGRPGPVLDLPFDVQVAEIEFDPDMYEPLPVYKPAASRMQIEKAVEMLI
130      140      150      160      170      180 X 190      200

QAERPVIIVAGGGVINADAAALLQFAELTSVPVI
      210      220      230
```

11. US-08-249-182-2 (1-6)

S03208 type III site-specific deoxyribonuclease (EC 3.1.2

ENTRY S03208 #type complete
TITLE type III site-specific deoxyribonuclease (EC 3.1.21.5) EcoP15
chain mod - Escherichia coli
ALTERNATE_NAMES type III restriction enzyme EcoP15 chain mod
ORGANISM #formal_name Escherichia coli
DATE 07-Jun-1990 #sequence_revision 31-Dec-1991 #text_change
28-Apr-1993
ACCESSIONS S03208
REFERENCE S01351
#authors Huembelin, M.; Suri, B.; Rao, D.N.; Hornby, D.P.; Eberle, H.;
Pripfl, T.; Kenel, S.; Bickle, T.A.
#journal J. Mol. Biol. (1988) 200:23-29
#title Type III DNA restriction and modification systems EcoP1 and
EcoP15. Nucleotide sequence of the EcoP1 operon, the EcoP15
mod gene and some EcoP1 mod mutants.
#cross-references MUID:88245189
#accession S03208
##molecule_type DNA
##residues 1-645 ##label HUE
##cross-references EMBL:X06288
##note the amino acid sequence from Fig. 6 is inconsistent with
the nucleotide sequence from Fig. 5 in having an
additional Met-Asp at the amino end

GENETICS

#gene mod

KEYWORDS hydrolase

SUMMARY #length 645 #molecular-weight 74221 #checksum 4238

SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                || ||
EQLISEWKSNIQVKNLLVNIGEEFASKYTGNELQEKYTQWFREHRSSELWPLDRYKYIDKGIYTGSSQSVHN
      260      270      280      290      300 X 310      320

PGKEGYRYDIIHPKTKKPCKQQLMGYRFPLDTM
      330      340      350
```

12. US-08-249-182-2 (1-6)

OYRTA1 guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 c

ENTRY OYRTA1 #type complete
 TITLE guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - rat
 ALTERNATE_NAMES guanylate cyclase, soluble, 77K chain
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 31-Dec-1993
 ACCESSIONS A38297
 REFERENCE A38297
 #authors Nakane, M.; Arai, K.; Saheki, S.; Kuno, T.; Buechler, W.; Murad, F.
 #journal J. Biol. Chem. (1990) 265:16841-16845
 #title Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase from rat lung.
 #cross-references MUID:91009100
 #accession A38297
 ##molecule_type mRNA
 ##residues 1-690 ##label NAK
 ##cross-references GB:M36075
 CLASSIFICATION #superfamily soluble guanylate cyclase; guanylate cyclase catalytic domain homology
 KEYWORDS cGMP synthesis; heterodimer; phosphorus-oxygen lyase
 FEATURE
 424-673 #domain guanylate cyclase catalytic domain #status predicted #label CAT
 SUMMARY #length 690 #molecular-weight 77566 #checksum 6429
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X   X
                                PLDVYK
                                |||||
GQIVQAKKFNEVTMLFSDIVGFTAICSGCSPLQVITMLNALYTRFDQCGELDVYKVETIGDAYCVAGGLHR
470      480      490      500      510      520 X      530

ESDTHAVQIALMALKMMELSNEVMSPHGEPKMR
540      550      560      570

```

13. US-08-249-182-2 (1-6)

OYB077 guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 c

ENTRY OYB077 #type complete
 TITLE guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - bovine
 ALTERNATE_NAMES guanylate cyclase, soluble, 77K chain
 ORGANISM #formal_name Bos primigenius taurus #common_name cattle
 DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 31-Dec-1993
 ACCESSIONS S10713; A38767
 REFERENCE S10713
 #authors Koesling, D.; Harteneck, C.; Humbert, P.; Bosserhoff, A.; Frank, R.; Schultz, G.; Boehne, E.
 #journal FEBS Lett. (1990) 266:128-132
 #title The primary structure of the larger subunit of soluble guanylyl cyclase from bovine lung. Homology between the two subunits of the enzyme.
 #cross-references MUID:90306336
 #accession S10713
 ##molecule_type mRNA
 ##residues 1-691 ##label KOE1
 ##cross-references ENBL:X54014
 #accession A38767

```

##molecule_type protein
##residues 118-133;226-232;286-293;319-330;412-417;557-571;629-637
##label KOE2
CLASSIFICATION #superfamily soluble guanylate cyclase; guanylate cyclase
catalytic domain homology
KEYWORDS cGMP synthesis; heterodimer; phosphorus-oxygen lyase
SUMMARY #length 691 #molecular-weight 77532 #checksum 2903
SEQUENCE

```

```

Initial Score = 5 Optimized Score = 5 Significance = 4.09
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

```

X X
PLDVYK
|||||
GHAVQAKRFGNVITLFSDIVGFTAICSQCSPLQVITMLNALYTRFDRCGELDVYKVETIGDAYCVAGGLHK
470 480 490 500 510 520 X 530 540

ESDTHAVQIALMALKMMELSHVVSPPHGEPIKMR
550 560 570

```

14. US-08-249-182-2 (1-6)

S23098 guanylate cyclase (EC 4.6.1.2), soluble, 81K chain

```

ENTRY S23098 #type complete
TITLE guanylate cyclase (EC 4.6.1.2), soluble, 81K chain - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
22-Nov-1993
ACCESSIONS S23098
REFERENCE S23097
#authors Giuli, G.; Scholl, U.; Bulle, F.; Guellaen, G.
#journal FEBS Lett. (1992) 304:83-88
#title Molecular cloning of the cDNAs coding for the two subunits of
soluble guanylyl cyclase from human brain.
#accession S23098
##status preliminary
##residues 1-717 ##label GIU
##cross-references EMBL:X66534
SUMMARY #length 717 #molecular-weight 81377 #checksum 6372
SEQUENCE

```

```

Initial Score = 5 Optimized Score = 5 Significance = 4.09
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

```

X X
PLDVYK
|||||
GQVVQAKKFSNVITLFSDIVGFTAICSQCSPLQVITMLNALYTRFDRCGELDVYKVETIAMPIVWLGLHK
470 480 490 500 510 X 520 530

ESDTHAVQIALMALKMMELSDVMSPPHGEPIKMR
540 550 560 570

```

15. US-08-249-182-2 (1-6)

A42163 Na+/myo-inositol cotransporter - dog

```

ENTRY A42163 #type complete
TITLE Na+/myo-inositol cotransporter - dog
ORGANISM #formal_name Canis lupus familiaris #common_name dog
DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change
03-May-1994

```

REFERENCE A42163
 #authors Kwon, H.M.; Yamauchi, A.; Uchida, S.; Preston, A.S.;
 Garcia-Perez, A.; Burg, M.B.; Handler, J.S.
 #journal J. Biol. Chem. (1992) 267:6297-6301
 #title Cloning of the cDNA for a Na(+)-myo-inositol cotransporter, a
 hypertonicity stress protein.
 #accession A42163
 ##status preliminary
 ##molecule_type mRNA
 ##residues 1-718 ##label KWD
 ##cross-references GB:M85068

SUMMARY #length 718 #molecular-weight 79545 #checksum 3177

SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.09
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 GSRAGCSNIAYPRLVMKLPVGLRGLMMAVMIAALMSDLDSIFNSASTIFTLDVYKLIRRSASSRELMIVGR
 350 360 370 380 390 X 400 410

IFVAFMVVISIAWVPIIVEMGGQMYLYIQEVAD
 420 430 440

> 0 <
 0| |0 IntelliGenetics
 > 0 <

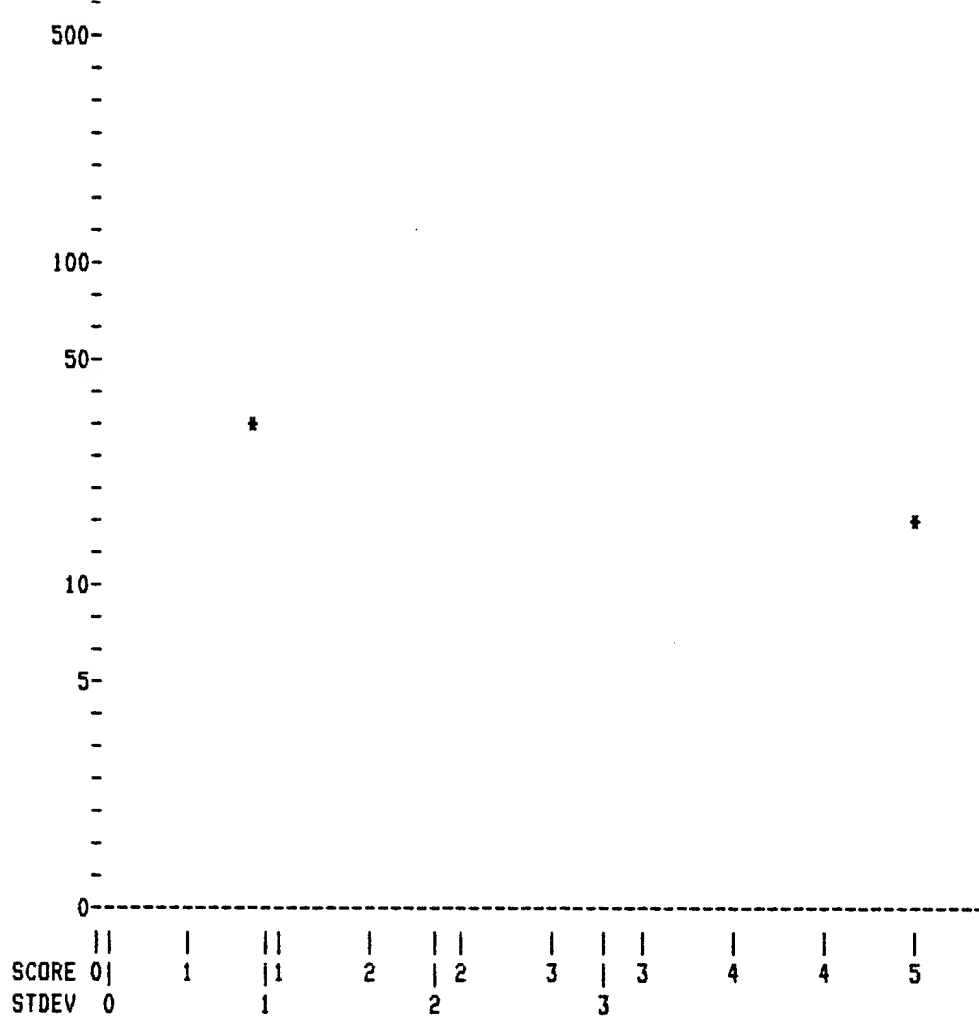
FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_2s.res made by on Thu 22 Sep 94 10:53:06-PDT.

Query sequence being compared:US-08-249-182-2 (1-6)
 Number of sequences searched: 36000
 Number of scores above cutoff: 4368

Results of the initial comparison of US-08-249-182-2 (1-6) with:
 Data bank : Swiss-Prot 28, all entries

100000-
 -
 N -
 U50000-
 M -
 B -
 E - *
 R -
 -
 O -
 F10000-
 - *
 S -
 E 5000*
 Q -
 U -
 E -
 N -
 C -
 E -
 S 1000-
 - *



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.86

Times:	CPU	Total Elapsed
	00:00:49.89	00:00:55.00

Number of residues:	12496420
Number of sequences searched:	36000
Number of scores above cutoff:	4368

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Score	Sig.	Frame
**** 4 standard deviations above mean ****							
1. YCX6_YEAST	HYPOTHETICAL 21.7 KD PROTEIN	190	5		5	4.63	0
2. ADT1_YEAST	ADP,ATP CARRIER PROTEIN 1 (AD	309	5		5	4.63	0
3. DRN1_STREQ	DEOXYRIBONUCLEASE PRECURSOR (327	5		5	4.63	0
4. EPIP_STAEP	EPIDERMIS PROCESSING SERINE P	461	5		5	4.63	0
5. CRT1_SYNP7	PHYTOENE DEHYDROGENASE (EC 1.	474	5		5	4.63	0
6. FUMA_ECOLI	FUMARATE HYDRATASE CLASS I, A	548	5		5	4.63	0
7. GCL_ECOLI	GLYOXYLATE CARBOXYLASE (EC 4.	592	5		5	4.63	0
8. T3M0_ECOLI	TYPE III RESTRICTION-MODIFICA	645	5		5	4.63	0
9. CYG3_RAT	GUANYLATE CYCLASE SOLUBLE, AL	690	5		5	4.63	0
10. CYG3_BOVIN	GUANYLATE CYCLASE SOLUBLE, AL	691	5		5	4.63	0
11. CYG5_HUMAN	GUANYLATE CYCLASE SOLUBLE, AL	717	5		5	4.63	0
12. NAMI_CANFA	SODIUM/MYO-INOSITOL COTRANSPO	718	5		5	4.63	0
13. DPOL_SULSO	DNA POLYMERASE (EC 2.7.7.7).	882	5		5	4.63	0
14. EBA1_PLAFC	ERYTHROCYTE-BINDING ANTIGEN E	1426	5		5	4.63	0
15. PR08_YEAST	PRE-mRNA SPLICING FACTOR PRP8	2413	5		5	4.63	0
**** 3 standard deviations above mean ****							
16. DESR_DESGI	DESULFOREDOXIN (DX).	36	4		4	3.47	0
17. APA1_COTJA	APOLIPOPROTEIN A-I (APO-AI) (36	4		4	3.47	0
18. LDH_PLAFA	L-LACTATE DEHYDROGENASE (EC 1	42	4		4	3.47	0
19. YCA3_BPT4	HYPOTHETICAL 5.1 KD PROTEIN I	45	4		4	3.47	0
20. DIUH_ACHDO	DIURETIC HORMONE (DH).	46	4		4	3.47	0
21. KCRB_SQUAC	CREATINE KINASE, B CHAIN (EC	52	4		4	3.47	0
22. SAS3_BACME	SMALL, ACID-SOLUBLE SPORE PRO	65	4		4	3.47	0
23. COX0_BOVIN	CYTOCHROME C OXIDASE POLYPEPT	70	4		4	3.47	0
24. PSAX_SYNY3	PHOTOSYSTEM I IRON-SULFUR CEN	80	4		4	3.47	0
25. PSAC_SYNP2	PHOTOSYSTEM I IRON-SULFUR CEN	80	4		4	3.47	0
26. PSAC_CHLRE	PHOTOSYSTEM I IRON-SULFUR CEN	80	4		4	3.47	0
27. YORY_PYRWO	HYPOTHETICAL 9.7 KD PROTEIN I	87	4		4	3.47	0
28. VFUS_ORFNZ	10 KD FUSION PROTEIN.	89	4		4	3.47	0
29. YPS1_SYNP2	HYPOTHETICAL 10.4 KD PROTEIN	92	4		4	3.47	0
30. CYC_TRYBB	CYTOCHROME C (FRAGMENTS).	93	4		4	3.47	0
31. CH10_ACYPS	10 KD CHAPERONIN (PROTEIN CPN	96	4		4	3.47	0
32. YXYB_CALSA	HYPOTHETICAL 10.7 KD PROTEIN	97	4		5	3.47	0
33. VG45_BPMLS	GENE 45 PROTEIN (GP45).	97	4		4	3.47	0
34. CH10_ECOLI	10 KD CHAPERONIN (PROTEIN CPN	97	4		4	3.47	0
35. ELIC_PHYCI	CINNAMOMIN.	98	4		4	3.47	0
36. CH10_BRUAB	10 KD CHAPERONIN (PROTEIN CPN	98	4		4	3.47	0
37. CH10_MYCBO	10 KD CHAPERONIN (PROTEIN CPN	99	4		4	3.47	0
38. CH10_MYCTU	10 KD CHAPERONIN (PROTEIN CPN	100	4		4	3.47	0
39. CH10_STRAL	10 KD CHAPERONIN (PROTEIN CPN	102	4		4	3.47	0
40. FINC_MOUSE	FIBRONECTIN (FN) (FRAGMENT).	103	4		4	3.47	0

1. US-08-249-182-2 (1-6)

YCX6_YEAST HYPOTHETICAL 21.7 KD PROTEIN IN TUP1-ABP1 INTERGEN

```

ID   YCX6_YEAST    STANDARD;       PRT;   190 AA.
AC   P25651;
DT   01-MAY-1992 (REL. 22, CREATED)
DT   01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT   01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 21.7 KD PROTEIN IN TUP1-ABP1 INTERGENIC REGION.
GN   YCR86W.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   DUSTERHOFT A., ERDMANN D., HEGEMANN J., PHILIPPSEN P., SCHWEITZER B.,

```

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; X59720; SCCHRIII.
 DR PIR; S19501; S19501.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 190 AA; 21751 MW; 199156 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 MDPLTVYKNSVKQIDSDADLLVANLVNENFVLSEKLDTKATEIKQLKQIDSLNAQVK
 X 10 20 30 40 50

2. US-08-249-182-2 (1-6)

ADT1_YEAST ADP,ATP CARRIER PROTEIN 1 (ADP/ATP TRANSLOCASE 1)

ID ADT1_YEAST STANDARD; PRT; 309 AA.
 AC P04710;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE ADP,ATP CARRIER PROTEIN 1 (ADP/ATP TRANSLOCASE 1) (ADENINE NUCLEOTIDE
 DE TRANSLOCATOR 1) (ANT 1).
 GN AAC1 OR PET9.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 87064348
 RA ADRIAN G.S., MCCAMMON M.T., MONTGOMERY D.L., DOUGLAS M.G.;
 RL MOL. CELL. BIOL. 6:626-634(1986).
 RN [2]
 RP SEQUENCE OF 265-309 FROM N.A.
 RC STRAIN=S288C;
 RM 91330299
 RA HOYT M.A., TOTIS L., ROBERTS B.T.;
 RL CELL 66:507-517(1991).
 CC -!- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
 CC MITOCHONDRIAL INNER MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- EACH CHAIN IS COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 DR EMBL; M12514; SCPET9.
 DR EMBL; M64706; SCBUB20.
 DR PIR; A24849; A24849.
 DR PROSITE; PS00215; MITOCH_CARRIER.
 KW MITOCHONDRION; INNER MEMBRANE; DUPLICATION; TRANSMEMBRANE; TRANSPORT;
 KW MULTIGENE FAMILY.
 FT TRANSMEM 17 34 1 (POTENTIAL).
 FT TRANSMEM 79 97 2 (POTENTIAL).
 FT TRANSMEM 122 139 3 (POTENTIAL).
 FT TRANSMEM 183 202 4 (POTENTIAL).
 FT TRANSMEM 222 239 5 (POTENTIAL).
 FT TRANSMEM 278 296 6 (POTENTIAL).
 SQ SEQUENCE 309 AA; 34120 MW; 502547 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
|||||

YAKWFAGNLFSGGAAGGLSLLFVYSLDYARTRLAADARGSKSTSQRQFNGLLDVYKKTCLKTDGLLGLYRGFV
120 130 140 150 160 X 170 180

PSVLGIIVYRGLYFGLYDSFKPVLLTGALEGSFV
190 200 210

3. US-08-249-182-2 (1-6)

DRN1_STREQ DEOXYRIBONUCLEASE PRECURSOR (EC 3.1.21.1) (STREPTO

ID DRN1_STREQ STANDARD; PRT; 327 AA.
AC P26295;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE DEOXYRIBONUCLEASE PRECURSOR (EC 3.1.21.1) (STREPTODORNASE) (DNASE).
GN SDC.
OS STREPTOCOCCUS EQUISIMILIS.
OC PROKARYOTA; FIRMICUTES; COCCI; STREPTOCOCCAEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RM 92039051
RA WOLINOWSKA R., CEGLOWSKI P., KOK J., VENEMA G.;
RL GENE 106:115-119(1991).
CC -!- FUNCTION: MAY HAVE A ROLE IN S.EQUISIMILIS VIRULENCE.
CC -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-
CC PHOSPHODINUCLEOTIDE AND 5'-PHOSPHOOLIGONUCLEOTIDE END-PRODUCTS.
DR EMBL; X17241; SESDC.
DR PIR; JT0584; JT0584.
KW ENDONUCLEASE; NUCLEASE; SIGNAL.
FT SIGNAL 1 24 OR 35 (POTENTIAL).
FT CHAIN 25 327 DEOXYRIBONUCLEASE.
SQ SEQUENCE 327 AA; 36844 MW; 583871 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||| ||

YGEYKDYYTVIGESNIDQSAFPKIYKTTTERVYKGQGTSEKRVTVSDVVYNPLDGYKRSTGAYGVVTKDMIDM
50 60 70 80 90 X 100 110

SKGYREKWETNPEPSGWFRFYNRADNEEISEKEY
120 130 140

4. US-08-249-182-2 (1-6)

EPIP_STAEP EPIDERMIN PROCESSING SERINE PROTEASE EPIP PRECURSO

ID EPIP_STAEP STANDARD; PRT; 461 AA.
AC P30199;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE EPIDERMIN PROCESSING SERINE PROTEASE EPIP PRECURSOR (EC 3.4.21.-).
GN EPIP.
OS STAPHYLOCOCCUS EPIDERMIDIS.
OG PLASMID PTU 32.
OC PROKARYOTA; FIRMICUTES; COCCI; MICROCOCCACEAE.

RP SEQUENCE FROM N.A.
 RC STRAIN=TU 3298 / DSM 3095;
 RM 92155237
 RA SCHNELL N., ENGELKE G., AUGUSTIN J., ROSENSTEIN R., UNGERMANN V.,
 RA GOETZ F., ENTIAN K.-D.;
 RL EUR. J. BIOCHEM. 204:57-68(1992).
 CC -!- FUNCTION: EPIP MIGHT BE THE PROTEASE WHICH CLEAVES THE MATURED
 CC LANTIBIOTIC FROM THE MODIFIED PREPEPTIDE.
 CC -!- SIMILARITY: BELONGS TO THE SUBTILASE PROTEASES FAMILY.
 DR EMBL; X62386; SEEPIDNA.
 DR PIR; S23420; S23420.
 DR PROSITE; PS00136; SUBTILASE_ASP.
 DR PROSITE; PS00137; SUBTILASE_HIS.
 DR PROSITE; PS00138; SUBTILASE_SER.
 KW HYDROLASE; SERINE PROTEASE; SIGNAL.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN ? 461 EPIDERMIN PROCESSING SERINE PROTEASE
 FT EPIP.
 FT ACT_SITE 149 149 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 402 402 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 461 AA; 51814 MW; 1103690 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||||
 SLAAPKVSGALALEIDKYQLKDQETAIELFKKKGIEKEKYMDKKHYGNGLDVYKLLKE
 410 420 430 440 450 X 460

5. US-08-249-182-2 (1-6)

CRTI_SYNP7 PHYTOENE DEHYDROGENASE (EC 1.3.-.-) (PHYTOENE DESA

ID CRTI_SYNP7 STANDARD; PRT; 474 AA.
 AC P26294;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE PHYTOENE DEHYDROGENASE (EC 1.3.-.-) (PHYTOENE DESATURASE).
 GN PDS.
 OS SYNECHOCOCCUS SP. (STRAIN PCC 7942) (ANACYSTIS NIDULANS R2).
 OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
 OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHRODOCCALES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91322511
 RA CHAMOVITZ D., PECKER I., HIRSCHBERG J.;
 RL PLANT MOL. BIOL. 16:967-974(1991).
 CC -!- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -!- PATHWAY: CAROTENOID BIOSYNTHESIS.
 CC -!- COFACTOR: NAD, NADP, OR FAD (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (PROBABLE).
 CC -!- ENZYME REGULATION: INHIBITED BY THE HERBICIDE NORFLURAZON IN A
 CC NON-COMPETITIVE WAY.
 CC -!- SIMILARITY: TO OTHER PLANTS OR CYANOBACTERIAL PHYTOENE
 CC DESATURASES.
 DR EMBL; X55289; SYPDSG.
 DR PIR; S16250; S16250.
 KW CAROTENOID BIOSYNTHESIS; OXIDOREDUCTASE; NAD; FAD;

FT NP_BIND 7 23 FAD OR NAD (POTENTIAL).
 FT VARIANT 403 403 V -> G (CONFERS RESISTANCE TO THE
 FT HERBICIDE NORFLURAZON; STRAIN NFZ4).
 SQ SEQUENCE 474 AA; 53296 MW; 1092205 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 || ||
 ELVFAPAKDWIGRSDDEDILAATMAEIEKLFPGHFSGENPARLRKYKIVKTPLSVYKATPGRQQYRPDQASPI
 350 360 370 380 390 400 X 410 420
 ANFFLTGDYTMQRYLASMEGAVLSCKLTAQAIIA
 430 440 450

6. US-08-249-182-2 (1-6)

FUMA_ECOLI FUMARATE HYDRATASE CLASS I, AEROBIC (EC 4.2.1.2) (

ID FUMA_ECOLI STANDARD; PRT; 548 AA.
 AC P00923;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE FUMARATE HYDRATASE CLASS I, AEROBIC (EC 4.2.1.2) (FUMARASE).
 GN FUMA.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 84221385
 RA MILES J.S., GUEST J.R.;
 RL NUCLEIC ACIDS RES. 12:3631-3642(1984).
 RN [2]
 RP IDENTIFICATION OF THE STRUCTURAL GENE.
 RM 86142617
 RA GUEST J.R., MILES J.S., ROBERTS R.E., WOODS S.A.;
 RL J. GEN. MICROBIOL. 131:2971-2984(1985).
 RN [3]
 RP BIOCHEMICAL ANALYSIS OF FUMA AND FUMC.
 RM 88193096
 RA WOODS S.A., SHWARTZBACH S.D., GUEST J.R.;
 RL BIOCHIM. BIOPHYS. ACTA 954:14-26(1988).
 RN [4]
 RP IRON-SULFUR CLUSTER.
 RA FLINT D.H., EMPTAGE M.H., GUEST J.R.;
 RL J. INORG. BIOCHEM. 36:306-306(1989).
 CC -!- CATALYTIC ACTIVITY: L-MALATE = FUMARATE + H(2)O.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- PATHWAY: TRICARBOXYLIC ACID CYCLE.
 CC -!- COFACTOR: IRON-SULFUR 4FE-4S CLUSTER.
 CC -!- ENZYME REGULATION: SUBJECT TO AEROBIC RESPIRATORY CONTROL AND
 CC CATABOLITE REPRESSION.
 CC -!- FUMA ACCOUNTS FOR ABOUT 80% OF THE FUMARASE ACTIVITY WHEN
 CC E.COLI GROWS AEROBICALLY.
 CC -!- THIS IS A CLASS-I FUMARASE. CLASS-I FUMARASES ARE THERMOLABILE.
 CC -!- SIMILARITY: 79% IDENTITY WITH E.COLI FUMB.
 DR EMBL; X00522; ECFUMA.
 DR PIR; A03531; UFECAQ.
 DR ECGENE; EG10356; FUMA.
 DR PROSITE; PS00163; FUMARATE_LYASES.

RW LYASE; TRICARBOXYLIC ACID CYCLE; IRON-SULFUR (4FE-4S).
FT METAL 318 318 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT ACT_SITE 397 397 POTENTIAL.
FT BINDING 463 463 CARBOXYL GROUP (POTENTIAL).
SQ SEQUENCE 548 AA; 60298 MW; 1466324 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||| ||
VLPTCQDTGTATIIVGKKGRVWTGGGDEAALARGVYNTYIEDNLRYSQAPLDMYKEVNTGTNLPAGIDLYA
110 120 130 140 150 X 160 170

VDGDEYKFLCIAKGGGSANKTYLYQETKALLTPG
180 190 200

7. US-08-249-182-2 (1-6)

GCL_ECOLI GLYOXYLATE CARBOLIGASE (EC 4.1.1.47) (TARTRONATE-S

ID GCL_ECOLI STANDARD; PRT; 592 AA.
AC P30146;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE GLYOXYLATE CARBOLIGASE (EC 4.1.1.47) (TARTRONATE-SEMIALDEHYDE
DE SYNTHASE).
GN GCL.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.
RC STRAIN=K12;
RM 93179387
RA CHANG Y.-Y., WANG A.-Y., CRONAN J.E. JR.;
RL J. BIOL. CHEM. 268:3911-3919(1993).
CC -!- CATALYTIC ACTIVITY: 2 GLYOXYLATE = TARTRONATE SEMIALDEHYDE +
CC CO(2).
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE FLAVOPROTEIN, AND MAGNESIUM ION.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- PATHWAY: GLYOXYLATE CATABOLISM.
CC -!- INDUCTION: BY GLYOXYLATE.
CC -!- SIMILARITY: WITH OTHER ENZYMES WHICH REQUIRE TPP.
DR EMBL; L03845; ECGCL.
DR PIR; JT0742; JT0742.
DR ECOGENE; EG11583; GCL.
KW LYASE; FLAVOPROTEIN; THIAMINE PYROPHOSPHATE; FAD; MAGNESIUM.
FT INIT_MET 0 0
SQ SEQUENCE 592 AA; 64600 MW; 1701474 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PLDVYK
||| ||
AVTVREAALVPRVLQQAFHLMRSGRPGPVLVDLPFDVQVAEIEFDPDMYEPLPVYKPAASRMQIEKAVEMLI
130 140 150 160 170 180 X 190 200

QAERPVIAGGGVINADAAALLQQAELTSVPVI
210 220 230

8. US-08-249-182-2 (1-6)

T3MO_ECOLI TYPE III RESTRICTION-MODIFICATION SYSTEM ECOP15 EN

ID T3MO_ECOLI STANDARD; PRT; 645 AA.
 AC P12364;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE TYPE III RESTRICTION-MODIFICATION SYSTEM ECOP15 ENZYME MOD
 DE (EC 2.1.1.72).
 GN MOD.
 OS ESCHERICHIA COLI.
 OG PLASMID P15B.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 88245189
 RA HUEMBELIN M., SURI B., RAO D.N., HORNBV D.P., EBERLE H., PRIPFL T.,
 RA KENEL S., BICKLE T.A.;
 RL J. MOL. BIOL. 200:23-29(1988).
 CC -!- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR BINDING THE SYSTEM-
 CC SPECIFIC DNA RECOGNITION SITE IN BOTH RESTRICTION AND THE
 CC METHYLATION OF ONE OF THE ADENOSYL RESIDUES IN THE RECOGNITION
 CC SITE 5'-CAGCAC-3'.
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA ADENINE =
 CC S-ADENOSYL-L-HOMOCYSTEINE + DNA 6-METHYLAMINOPURINE.
 CC -!- SUBUNIT: TYPE III RESTRICTION SYSTEM CONTAINS TWO DIFFERENT
 CC SUBUNITS: THE PRODUCTS OF THE RES AND MOD GENES.
 CC -!- SUBUNIT: TYPE III MODIFICATION ENZYME IS A TETRAMER OF A
 CC SINGLE SUBUNIT, THE PRODUCT OF THE MOD GENE.
 DR EMBL; X06288; ECP15BMD.
 DR PIR; S03208; S03208.
 DR REBASE; ECOP15I; RELEASE 9402.
 DR PROSITE; PS00092; N6_MTASE.
 KW TRANSFERASE; METHYLTRANSFERASE; DNA-BINDING; RESTRICTION SYSTEM;
 KW PLASMID.
 SQ SEQUENCE 645 AA; 74222 MW; 2156590 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||| ||

EQLISEWKSNI SDVKNLLVNIGEEFASKYTCNELQEKYTQWFREHRSELWPLDRYKYIDKDGIIYTGSGQSVHN
 260 270 280 290 300 X 310 320

PGKEGYRYDIIHPKTKKPKCKQQLMGYRFPLDTM
 330 340 350

9. US-08-249-182-2 (1-6)

CYG3_RAT GUANYLATE CYCLASE SOLUBLE, ALPHA-1 CHAIN (EC 4.6.1

ID CYG3_RAT STANDARD; PRT; 690 AA.
 AC P19686;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE GUANYLATE CYCLASE SOLUBLE, ALPHA-1 CHAIN (EC 4.6.1.2) (82 KD CHAIN).
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=LUNG;
 RM 91009100
 RA NAKANE M., ARAI K., SAHEKI S., KUND T., BUECHLER W., MURAD F.;
 RL J. BIOL. CHEM. 265:16841-16845(1990).
 CC -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE FORMS AND
 CC MEMBRANE-ASSOCIATED RECEPTOR FORMS.
 CC -!- SIMILARITY: TO OTHER GUANYLATE CYCLASES.
 DR EMBL; M57405; RNGCSB.
 DR PIR; A38297; DYRTA1.
 DR PROSITE; PS00452; GUANYLATE_CYCLASES.
 KW LYASE; CGMP SYNTHESIS; MULTIGENE FAMILY.
 SQ SEQUENCE 690 AA; 77566 MW; 2459503 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 GQIVQAKKFNEVTMLFSDIVGFTAICSQCSPLQVITMLNALYTRFDQCGELDVYKVETIGDAYCVAGGLHR
 470 480 490 500 510 520 X 530
 ESDTHAVQIALMALKMMELSNEVNSPHGEPIKMR
 540 550 560 570

10. US-08-249-182-2 (1-6)
 CYG3_BOVIN GUANYLATE CYCLASE SOLUBLE, ALPHA-1 CHAIN (EC 4.6.1

ID CYG3_BOVIN STANDARD; PRT; 691 AA.
 AC P19687;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE GUANYLATE CYCLASE SOLUBLE, ALPHA-1 CHAIN (EC 4.6.1.2) (73 KD CHAIN).
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ARTIODACTYLA.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=ADRENAL MEDULLA;
 RM 90306336
 RA KOESLING D., HARTENECK C., HUMBERT P., BOSSERHOFF A., FRANK R.,
 RA SCHULTZ G., BOEHME E.;
 RL FEBS LETT. 266:128-132(1990).
 CC -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE FORMS AND
 CC MEMBRANE-ASSOCIATED RECEPTOR FORMS.
 CC -!- SIMILARITY: TO OTHER GUANYLATE CYCLASES.
 DR EMBL; X54014; BTGUCY73.
 DR PIR; S10713; OYB077.
 DR PROSITE; PS00452; GUANYLATE_CYCLASES.
 KW LYASE; CGMP SYNTHESIS; MULTIGENE FAMILY.
 SQ SEQUENCE 691 AA; 77532 MW; 2500456 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1

Gaps
 X X
 PLDVYK
 |||||
 GHAVQAKRFGNVTMLFSDIVGFTAICSQCSPLQVITMLNLYTRFDRCGELDVYKVETIGDAYCVAGGLHK
 470 480 490 500 510 520 X 530 540
 ESDTHAVQIALMALKMMELSHEVVSPHGEPIKMR
 550 560 570

11. US-08-249-182-2 (1-6)
 CYG5_HUMAN GUANYLATE CYCLASE SOLUBLE, ALPHA-3 CHAIN (EC 4.6.1

ID CYG5_HUMAN STANDARD; PRT; 717 AA.
 AC 002108;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE GUANYLATE CYCLASE SOLUBLE, ALPHA-3 CHAIN (EC 4.6.1.2).
 GN GUCSA3.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RM 92316204
 RA CIVILI G., SCHOLL U., BULLE F., GUELLAEEN G.;
 RL FEBS LETT. 304:83-88(1992).
 CC -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE FORMS AND
 CC MEMBRANE-ASSOCIATED RECEPTOR FORMS.
 CC -!- SIMILARITY: TO OTHER GUANYLATE CYCLASES.
 DR EMBL; X66534; HSGCSAA.
 DR PIR; S23098; S23098.
 DR NIM; 139396; TENTH EDITION.
 DR PROSITE; PS00452; GUANYLATE_CYCLASES.
 KW LYASE; CGMP SYNTHESIS; MULTIGENE FAMILY.
 SQ SEQUENCE 717 AA; 81377 MW; 2740969 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 |||||
 GQVVQAKKFSNVTMLFSDIVGFTAICSQCSPLQVITMLNLYTRFDRCGELDVYKVETIANPIVWLGLHK
 470 480 490 500 510 X 520 530
 ESDTHAVQIALMALKMMELSDEVNSPHGEPIKMR
 540 550 560 570

12. US-08-249-182-2 (1-6)
 NAMI_CANFA SODIUM/MYO-INOSITOL COTRANSPORTER (NA(+)/MYO-INOSI

ID NAMI_CANFA STANDARD; PRT; 718 AA.
 AC P31637;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)

DE SODIUM/HYD-INOITOL COTRANSPORTER (NAI777HYD-INOITOL COTRANSPORTER).
 GN SMIT.
 OS CANIS FAMILIARIS (DOG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; CARNIVORA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RM 92210609
 RA KWON H.M., YAMAUCHI A., UCHIDA S., PRESTON A.S., GARCIA-PEREZ A.,
 RA BURG M.B., HANDLER J.S.;
 RL J. BIOL. CHEM. 267:6297-6301(1992).
 CC -!- FUNCTION: PREVENTS INTRACELLULAR ACCUMULATION OF HIGH
 CC CONCENTRATIONS OF MYO-INOITOL (AN OSMOLYTE) THAT RESULT IN
 CC IMPAIRMENT OF CELLULAR FUNCTION.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: BRAIN AND KIDNEY.
 CC -!- INDUCTION: MEDIUM HYPERTONICITY.
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).
 DR EMBL; M85068; CFSMIT.
 DR PROSITE; PS00456; NA_SOLUT_SYMP_1.
 DR PROSITE; PS00457; NA_SOLUT_SYMP_2.
 KW TRANSPORT; TRANSMEMBRANE; SODIUM TRANSPORT; SYMPORT; GLYCOPROTEIN.
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 10 29 POTENTIAL.
 FT DOMAIN 30 38 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 39 57 POTENTIAL.
 FT DOMAIN 58 86 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 87 110 POTENTIAL.
 FT DOMAIN 111 123 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 124 144 POTENTIAL.
 FT DOMAIN 145 157 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 158 183 POTENTIAL.
 FT DOMAIN 184 186 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 187 205 POTENTIAL.
 FT DOMAIN 206 303 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 304 324 POTENTIAL.
 FT DOMAIN 325 353 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 354 376 POTENTIAL.
 FT DOMAIN 377 406 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 407 430 POTENTIAL.
 FT DOMAIN 431 443 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 444 462 POTENTIAL.
 FT DOMAIN 463 510 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 511 532 POTENTIAL.
 FT DOMAIN 533 695 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 696 716 POTENTIAL.
 FT CARBOHYD 32 32 POTENTIAL.
 FT SITE 24 24 IMPLICATED IN SODIUM COUPLING
 FT (BY SIMILARITY).
 FT SITE 285 285 IMPLICATED IN SODIUM COUPLING
 FT (BY SIMILARITY).
 SQ SEQUENCE 718 AA; 79545 MW; 2760723 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PLDVYK
 ||||
 GSRA GCSN IAYPRL VMKLV PVGLR GLMMA VMIAALMSDLSIFNSASTIFTLDVYKLIRRSASSRELMIVGR
 350 360 370 380 390 X 400 410
 IFVAFM VVISIA MWPII VEMOGGQMYLYIQEVAD
 420 430 440

13. US-08-249-182-2 (1-6)

DPOL_SULSO DNA POLYMERASE (EC 2.7.7.7).

ID DPOL_SULSO STANDARD; PRT; 882 AA.
 AC P26811;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE DNA POLYMERASE (EC 2.7.7.7).
 GN POLS.
 OS SULFOLOBUS SOLFATARICUS.
 OC PROKARYOTA; MENDOSICUTES; ARCHAEABACTERIA; SULFOLOBALES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT 4;
 RM 92310966
 RA PISANI F.M., MARTINO C., ROSSI M.;
 RL NUCLEIC ACIDS RES. 20:2711-2716(1992).
 CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
 CC N PYROPHOSPHATE + DNA(N).
 CC -!- SIMILARITY: BELONGS TO FAMILY B OF DNA POLYMERASES.
 DR EMBL; X64466; SSPOLS.
 DR PIR; S23019; S23019.
 DR PROSITE; PS00116; DNA_POLYMERASE_B.
 KW DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION; DNA-BINDING.
 SQ SEQUENCE 882 AA; 101333 MW; 4258623 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                PLDVYK
                                ||| ||
VKELMISINSPNDVKEIKRKIVDVVKGSYEKLKNKGYNLDELAFKVHLSKPLDAYKKNTPGHVKAALQLRPF
  740      750      760      770      780  X  790      800

GYNVLPREDIIYVVKVRSKDGVPVQLAKVTEIDA
  810      820      830

```

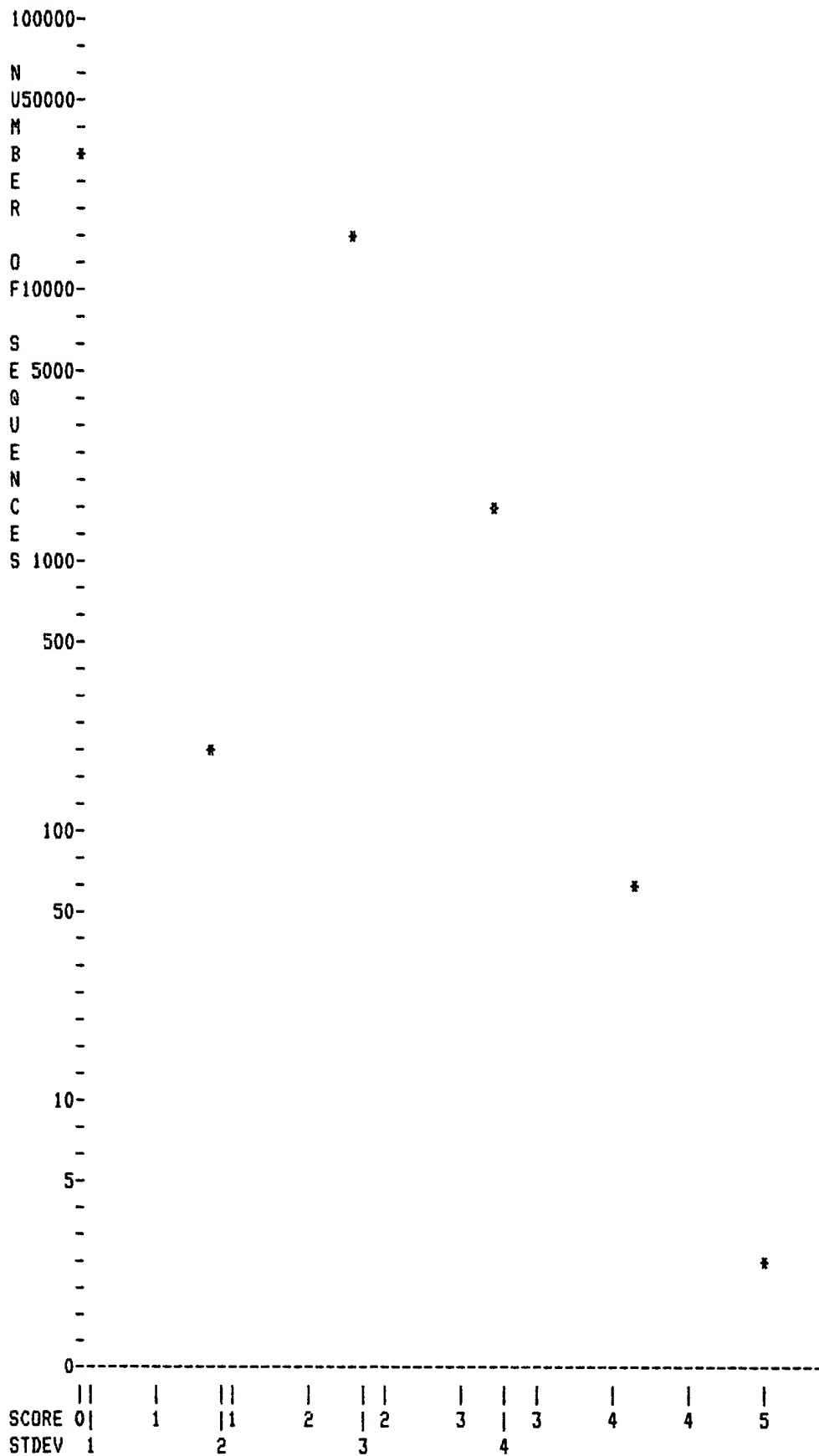
14. US-08-249-182-2 (1-6)

EBA1_PLAFC ERYTHROCYTE-BINDING ANTIGEN EBA-175.

ID EBA1_PLAFC STANDARD; PRT; 1426 AA.
 AC P19214;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-AUG-1991 (REL. 19, LAST ANNOTATION UPDATE)
 DE ERYTHROCYTE-BINDING ANTIGEN EBA-175.
 OS PLASMODIUM FALCIPARUM (ISOLATE CAMP / MALAYSIA).
 OC EUKARYOTA; PROTOZOA; APICOMPLEXA; SPOROZOA; COCCIDIA; EUCCOCCIDIIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LATE SCHIZONT;
 RM 90377299
 RA SIM B.K.L.;
 RL MOL. BIOCHEM. PARASITOL. 41:293-296(1990).
 DR EMBL; X52524; PFEBA175.
 DR PIR; S11561; S11561.
 KW ANTIGEN.
 FT DOMAIN 159 1101 ESSENTIAL FOR BINDING TO
 FT ERYTHROCYTES.
 FT VARIANT 1028 1028 E -> V (IN STRAINS FCR3 AND ITG).

Query sequence being compared:US-08-249-182-3 (1-5)
 Number of sequences searched: 42145
 Number of scores above cutoff: 4652

Results of the initial comparison of US-08-249-182-3 (1-5) with:
 Data bank : A-GeneSeq 15, all entries



PARAMETERS

SEQUENCE 1426 AA; 168155 MW; 10311930 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

          X   X
          PLDVYK
          |||||
MKCNISIVFFASFFVLVFAKARNEYDIKENEKFLDVYKEKFNELDKKKGNGVQKTDKKIFTFIENKLDILNN
      10      20      30 X      40      50      60      70
```

SKFNKRWKSYPGTPDNI
80

15. US-08-249-182-2 (1-6)

PRO8_YEAST PRE-MRNA SPLICING FACTOR PRP8.

ID PRO8_YEAST STANDARD; PRT; 2413 AA.
AC P33334;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE PRE-MRNA SPLICING FACTOR PRP8.
GN PRP8 OR RNA8.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA HODGES P.E., JACKSON S.P., BROWN J.D., BEGGS J.D.;
RL SUBMITTED (JUL-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP CHARACTERIZATION.
RM 88216580
RA JACKSON S.P., LOSSKY M., BEGGS J.D.;
RL MOL. CELL. BIOL. 8:1067-1075(1988).
CC -!- FUNCTION: INVOLVED IN PRE-MRNA SPLICING. U5 SNRNP PROTEIN.
CC APPEARS TO CONTACT THE PRE-MRNA DURING SPLICING.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: TO C.ELEGANS PROTEIN C50C3.6.
DR EMBL; 224732; SCPRP8GNA.
DR PIR; S34670; S34670.
KW MRNA PROCESSING; MRNA SPLICING; SPLICEOSOME; NUCLEAR PROTEIN.
SQ SEQUENCE 2413 AA; 279501 MW; 19891968 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.63
Residue Identity = 83% Matches = 5 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

          X   X
          PLDVYK
          |||||
MENYQNISPVYSVDPLEKITDAYLDQYLWYEADQKLFNPWIKPSDSEIPPLLVYKWTQGINNLSEIWDVSR
      980      990      1000      1010      1020 X      1030      1040

GQSAVLLETTLGEMAEKIDFTLLNRLLRLIVDPN
      1050      1060      1070
```

> 0 <
0| 10 IntelliGenetics
> 0 <

Seq 3

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_3a.res made by on Thu 22 Sep 94 10:29:31-PDT.

#title The DNA-binding protein hup from mesophilic and thermophilic
Bacilli: Gene cloning, overproduction and purification.

#cross-references MUID:92354934

#accession JC1206

##molecule_type DNA

##residues 1-90 ##label PAD

##cross-references GB:M73501

COMMENT This protein binds to both single-stranded and double-stranded DNA
and to RNA in stoichiometric amounts.

GENETICS

#gene hup

CLASSIFICATION #superfamily bacterial DNA-binding protein

KEYWORDS DNA binding; RNA binding

SUMMARY #length 90 #molecular-weight 9716 #checksum 4525

SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.06
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X

YPAFK

||||

DAVFDSITEALRKGDQVQLIGFGNFVRRERAARKGRNPQTGEEMEIPASKVPAFKPGKALKDAVK

30 40 50 60 70 X 80 90

> 0 <

0| |0 IntelliGenetics

> 0 <

FastDB - Fast Pairwise Comparison of Sequences

Release 5.4

Results file u249_3s.res made by on Thu 22 Sep 94 10:58:04-PDT.

Query sequence being compared:US-08-249-182-3 (1-5)

Number of sequences searched: 36000

Number of scores above cutoff: 3832

Results of the initial comparison of US-08-249-182-3 (1-5) with:
Data bank : Swiss-Prot 28, all entries

100000-

-

N -

U50000-

M -

B -

*

E -

R -

-

D -

F10000-

*

S -

E 5000-

Q -

*

U -

E -

N -

C -

E -

S 1000-

-

-

500-

KAEKGW

80

14. US-08-249-182-3 (1-5)

CCAI6 cytochrome c6 - *Anabaena variabilis*

ENTRY CCAI6 #type complete
TITLE cytochrome c6 - *Anabaena variabilis*
ALTERNATE_NAMES cytochrome c553; soluble cytochrome f
ORGANISM #formal_name *Anabaena variabilis*
DATE 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change
19-May-1994
ACCESSIONS A94488; A93182; A00105
REFERENCE A94488
#authors Beecher, J.; Margoliash, E.
#citation unpublished results, cited by Ulrich, E.L., Krogmann, D.W.,
and Markley, J.L., J. Biol. Chem. 257, 9356-9364, 1982
#accession A94488
##molecule_type protein
##residues 1-86 ##label BEE
REFERENCE A93182
#authors Aitken, A.
#journal Nature (1976) 263:793-796
#title Protein evolution in cyanobacteria.
#cross-references MUID:77056395
#accession A93182
##molecule_type protein
##residues 1-22;30-39;56-81,'D',83,'D',85-86 ##label AIT
CLASSIFICATION #superfamily cytochrome c; cytochrome c homology
KEYWORDS electron transfer; heme; photosynthesis
FEATURE
14,17 #binding_site heme (Cys) (covalent) #status predicted\
18,58 #binding_site heme iron (His, Met) (axial ligands)
#status predicted
SUMMARY #length 86 #molecular-weight 8973 #checksum 5450
SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.06
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X

YPAFK

||||

KIFSANCASCHAGGKNLGVAQKTLKKADLEKYGAYSAMAIGAQVTNGKNAMPFAFKGRLEKPEEIZBVAAYVLG

10 20 30 40 50 60 X 70

KAEAEWK

80

15. US-08-249-182-3 (1-5)

JC1206 DNA-binding protein HU - *Bacillus caldotenax*

ENTRY JC1206 #type complete
TITLE DNA-binding protein HU - *Bacillus caldotenax*
ORGANISM #formal_name *Bacillus caldotenax*
DATE 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change
23-Mar-1993
ACCESSIONS JC1206
REFERENCE JC1205
#authors Padas, P.M.; Wilson, K.S.; Vorgias, C.E.
#journal Gene (1992) 117:39-44

```

#accession A00093
##molecule_type protein
##residues 1-82 ##label AMB
GENETICS
#gene nirM
CLASSIFICATION #superfamily cytochrome c; cytochrome c homology
KEYWORDS electron transfer; heme; oxidative phosphorylation
FEATURE
  12,15 #binding_site heme (Cys) (covalent) #status predicted\
  16,61 #binding_site heme iron (His, Met) (axial ligands)
        #status predicted
SUMMARY #length 82 #molecular-weight 8612 #checksum 6641
SEQUENCE

```

```

Initial Score      =      4  Optimized Score =      4  Significance =  3.06
Residue Identity   =    80%  Matches          =      4  Mismatches   =      1
Gaps               =      0  Conservative Substitutions =      0

```

```

          X  X
          YPAFK
          ||||
GDGEALFKSKPCAACHSIDAKLVGPAFKEVAAKYAGGDGAADLLAGHIKNGSOGVWGPIPMPPNPVTEEEAK
      10      20  X   30      40      50      60      70

ILAEWI

```

13. US-08-249-182-3 (1-5)

CCPB6 cytochrome c6 - *Plectonema boryanum*

```

ENTRY      CCPB6      #type complete
TITLE      cytochrome c6 - Plectonema boryanum
ALTERNATE_NAMES cytochrome c553; soluble cytochrome f
ORGANISM    #formal_name Plectonema boryanum
DATE        #sequence_revision 07-May-1981 #text_change 19-May-1994
ACCESSIONS  A00109
REFERENCE   A00109
#authors    Aitken, A.
#journal     Eur. J. Biochem. (1977) 78:273-279
#title       Purification and primary structure of cytochrome f from the
              cyanobacterium, Plectonema boryanum.
#cross-references MUID:78023897
#contents    CCAP 1462/2
#accession   A00109
##molecule_type protein
##residues   1-85 ##label AIT
COMMENT     Plectonema is a genus of filamentous blue-green algae.
CLASSIFICATION #superfamily cytochrome c; cytochrome c homology
KEYWORDS      electron transfer; heme; photosynthesis
FEATURE
  14,17 #binding_site heme (Cys) (covalent) #status
        experimental\
  18,58 #binding_site heme iron (His, Met) (axial ligands)
        #status predicted
SUMMARY #length 85 #molecular-weight 8576 #checksum 3761
SEQUENCE

```

```

Initial Score      =      4  Optimized Score =      4  Significance =  3.06
Residue Identity   =    80%  Matches          =      4  Mismatches   =      1
Gaps               =      0  Conservative Substitutions =      0

```

```

          X  X
          YPAFK
          ||||
KVFNANCAACHASGGG@INCAKTLKKNALTANGKDTVEAIVAQVTNCKGAMPAFKRLSDDQIQSVALYVLD

```


#submission submitted to the EMBL Data Library, July 1993

#accession S35946

##status preliminary

##residues 1-43 ##label BAC

##cross-references EMBL:X74279

SUMMARY #length 43 #checksum 2542

SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.06
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X

YPAFK

||||

NLNDMRQFAEMHDIEITVIDNDTRLPAFKDALRWNEVYYGFRR

10 20 X 30 40

11. US-08-249-182-3 (1-5)

S10561 photosystem II 21K protein - barley

ENTRY S10561 #type complete

TITLE photosystem II 21K protein - barley

ORGANISM #formal_name Hordeum vulgare #common_name barley

DATE 02-Dec-1993; #sequence_revision 02-Dec-1993; #text_change
02-Dec-1993

ACCESSIONS S10561

REFERENCE S10561

#authors Morishige, D.T.; Anandan, S.; Jaing, J.T.; Thornber, J.P.

#journal FEBS Lett. (1990) 264:239-242

#title Amino-terminal sequence of the 21 kDa apoprotein of a minor
light-harvesting pigment-protein complex of the photosystem
II antenna (LHC II_d/CP 24).

#cross-references MUID:90292215

#accession S10561

##status preliminary

##residues 1-66 ##label MOR

SUMMARY #length 66 #molecular-weight 7169 #checksum 8176

SEQUENCE

Initial Score = 4 Optimized Score = 4 Significance = 3.06
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X

YPAFK

||||

AAAGKKSWIPAFKSDAEFINPSWLDGSLPGDFGFDPLGLGKDPFLKHWYGWAELIXXWAMXA

10 X 20 30 40 50 60

12. US-08-249-182-3 (1-5)

CCPS55 cytochrome c551 - Pseudomonas stutzeri (strain 221)

ENTRY CCPS55 #type complete

TITLE cytochrome c551 - Pseudomonas stutzeri (strain 221)

ORGANISM #formal_name Pseudomonas stutzeri

DATE #sequence_revision 13-Jul-1981 #text_change 19-May-1994

ACCESSIONS A00093

REFERENCE A90266

#authors Ambler, R.P.; Wynn, M.

#journal Biochem. J. (1973) 131:485-498

#title The amino acid sequences of cytochromes c-551 from three
species of Pseudomonas.

#cross-references MUID:73224976

```

#accession      A42327
##status        preliminary
##molecule_type protein
##residues      1-114 ##label STR
##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
                  NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
                  NCBIP:78509; NCBIP:78508; NCBIP:78503
##note          sequence extracted from NCBI backbone
SUMMARY         #length 114 #checksum 7335
SEQUENCE

```

```

Initial Score    =      5  Optimized Score =      5  Significance =  4.08
Residue Identity = 100%  Matches           =      5  Mismatches  =      0
Gaps             =      0  Conservative Substitutions =      0

```

```

                                X  X
                                YPAFK
                                ||||
GPTFKGGQPLWITATKSPPFENINLYYDVPWNETIPEEVTXPNYLQAEVSYPAFKPXLDVYKWHVAAN
  50          60          70          80          90         100        110

```

9. US-08-249-182-3 (1-5)

S12888 DNA-binding protein II - *Thermus aquaticus*

```

ENTRY           S12888    #type complete
TITLE           DNA-binding protein II - Thermus aquaticus
ORGANISM        #formal_name Thermus aquaticus
DATE            21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change
                21-Nov-1993
ACCESSIONS      S12888
REFERENCE       S12888
#authors        Zierer, R.; Choli, D.
#journal        FEBS Lett. (1990) 273:59-62
#title          The primary structure of DNA binding protein II from the
                extreme thermophilic bacterium Thermus thermophilus.
#cross-references MUID:91032203
#accession      S12888
##status        preliminary
##residues      1-95 ##label ZIE
SUMMARY         #length 95 #molecular-weight 10163 #checksum 68
SEQUENCE

```

```

Initial Score    =      5  Optimized Score =      5  Significance =  4.08
Residue Identity = 100%  Matches           =      5  Mismatches  =      0
Gaps             =      0  Conservative Substitutions =      0

```

```

                                X  X
                                YPAFK
                                ||||
DALLAKVEEALANGSKVQLTGFGTFEVRKRKARTGVKPGTKEKIKIPATQYPAFKPGKALKDKVK
  40          50          60          70          80         X    90

```

10. US-08-249-182-3 (1-5)

S35946 araA protein - *Escherichia coli* (fragment)

```

ENTRY           S35946    #type fragment
TITLE           araA protein - Escherichia coli (fragment)
ORGANISM        #formal_name Escherichia coli
DATE            22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
                22-Nov-1993
ACCESSIONS      S35946
REFERENCE       S35946
#authors        Bachellier, S.; Saurin, W.; Perrin, D.; Hofnung, M.; Gilson,
                E.

```

Gaps - - - - - Conservative Substitutions - - - - -

```

                                X  X
                                YPAFK
                                |||||
SLRTEENLAVVKQIPTEKLLLETDAPWCEIKRTHASFQYLAKYQEVDFEYPAFKSVKKNKLADKLNAEELY
  310      320      330      340      350  X  360      370

MVKGRNEPCNMEQVAIVVSEVKDVLATLIDTT
  380      390      400

```

7. US-08-249-182-3 (1-5)

S39833 hypothetical protein YBL0511 - yeast (*Saccharomyces*

ENTRY S39833 #type complete
 TITLE hypothetical protein YBL0511 - yeast (*Saccharomyces cerevisiae*)
 ORGANISM #formal_name *Saccharomyces cerevisiae*
 DATE 27-May-1994; #sequence_revision 27-May-1994; #text_change 27-May-1994
 ACCESSIONS S39833
 REFERENCE S39824
 #authors Scherens, B.; el Bakkoury, M.; Vierendeels, F.; Dubois, E.; Messenguy, F.
 #journal Yeast (1993) 9:1355-1371
 #title Sequencing and functional analysis of a 32 560 bp segment on the left arm of yeast chromosome II. Identification of 26 open reading frames, including the KIP1 and SEC17 genes.
 #accession S39833
 ##status preliminary
 ##residues 1-418 ##label SCH
 ##cross-references EMBL:Z23261
 SUMMARY #length 418 #molecular-weight 47390 #checksum 9974
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.08
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                YPAFK
                                |||||
SLRTEENLAVVKQIPTEKLLLETDAPWCEIKRTHASFQYLAKYQEVDFEYPAFKSVKKNKLADKLNAEELY
  310      320      330      340      350  X  360      370

MVKGRNEPCNMEQVAIVVSEVKDVLATLIDTT
  380      390      400

```

8. US-08-249-182-3 (1-5)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name *Homo sapiens* #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337

MDARATIPENHIARTVILPQGYADDEVIYPAFKWL RDEQPLAMAHIEGYDPMWIATKHADVMQIGKQPLFSN

10 20 30 X 40 50 60 70

AEQSEILYDQ

80

5. US-08-249-182-3 (1-5)

S27653 cytochrome P450-terpredoxin - Pseudomonas sp.

ENTRY S27653 #type complete
 TITLE cytochrome P450-terpredoxin - Pseudomonas sp.
 ORGANISM #formal_name Pseudomonas sp.
 DATE 25-Feb-1994; #sequence_revision 25-Feb-1994; #text_change 25-Feb-1994
 ACCESSIONS S27653
 REFERENCE S27651
 #authors Peterson, J.A.; Lu, J.Y.; Geisselsoder, J.; Graham-Lorence, S.; Carmona, C.; Witney, F.; Lorence, M.C.
 #submission submitted to the EMBL Data Library, April 1992
 #accession S27653
 ##status preliminary
 ##residues 1-428 ##label PET
 ##cross-references EMBL:M91440
 SUMMARY #length 428 #molecular-weight 47922 #checksum 8787
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.08
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X

YPAFK

||||

MDARATIPENHIARTVILPQGYADDEVIYPAFKWL RDEQPLAMAHIEGYDPMWIATKHADVMQIGKQPLFSN

10 20 30 X 40 50 60 70

AEQSEILYDQ

80

6. US-08-249-182-3 (1-5)

S37334 hypothetical protein YBL0511 - yeast (Saccharomyce

ENTRY S37334 #type complete
 TITLE hypothetical protein YBL0511 - yeast (Saccharomyces cerevisiae)
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
 ACCESSIONS S37334
 REFERENCE S37325
 #authors Scherens, B.
 #submission submitted to the EMBL Data Library, July 1993
 #accession S37334
 ##molecule_type DNA
 ##residues 1-418 ##label SCH
 ##cross-references EMBL:Z23261
 GENETICS
 #map_position 2
 SUMMARY #length 418 #molecular-weight 47390 #checksum 9974
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.08
 Residue Identity = 100% Matches = 5 Mismatches = 0

ENTRY S24389 #type complete
 TITLE prismane protein - Desulfovibrio desulfuricans
 ORGANISM #formal_name Desulfovibrio desulfuricans
 DATE 25-Feb-1994; #sequence_revision 25-Feb-1994; #text_change 25-Feb-1994
 ACCESSIONS S24389
 REFERENCE S24389
 #authors Stokkermans, J.P.W.G.; van den Berg, W.A.M.; van Dongen, W.M.A.M.; Veeger, C.
 #journal Biochim. Biophys. Acta (1992) 1132:83-87
 #title The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from Desulfovibrio desulfuricans (ATCC 27774).
 #cross-references MUID:92379097
 #accession S24389
 ##status preliminary
 ##residues 1-545 ##label STD
 ##cross-references EMBL:Z11975
 SUMMARY #length 545 #molecular-weight 58659 #checksum 9858
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.08
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 EITKVNIGVGSNPGILISGHDRLDLEMLLKQTEGTGVDVYTHSEMLPAHYYPAFKKYAHFKGNYGNAWWKQK
 230 240 250 260 270 X X 280 290
 EEFESEFNGPVLTTNCLVPPKDSYKDRVYTTGI
 300 310 320

4. US-08-249-182-3 (1-5)

A42971 cytochrome P-450terp, P-450terp - Pseudomonas sp.

ENTRY A42971 #type complete
 TITLE cytochrome P-450terp, P-450terp - Pseudomonas sp.
 ORGANISM #formal_name Pseudomonas
 DATE 04-Mar-1993; #sequence_revision 04-Mar-1993; #text_change 04-Mar-1993
 ACCESSIONS A42971
 REFERENCE A42971
 #authors Peterson, J.A.; Lu, J.Y.; Geisselsoder, J.; Graham-Lorence, S.; Carnona, C.; Witney, F.; Lorence, M.C.
 #journal J. Biol. Chem. (1992) 267:14193-14203
 #title Cytochrome P-450terp. Isolation and purification of the protein and cloning and sequencing of its operon.
 #cross-references MUID:92332528
 #accession A42971
 ##status preliminary
 ##molecule_type nucleic acid; protein
 ##residues 1-428 ##label PET
 ##cross-references NCBIP:108469
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 428 #molecular-weight 47922 #checksum 8787
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.08
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK

```

##molecule_type DNA
##residues 1-1096 ##label KAT
GENETICS
#gene pula
KEYWORDS glycosidase; hydrolase
FEATURE
  1-19 #domain signal sequence #label SIG\
  20-1096 #protein alpha-dextrin endo-1,6-alpha-glucosidase #label
          MAT
SUMMARY #length 1096 #molecular-weight 119335 #checksum 1390
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.08
Residue Identity   =  100%  Matches          =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

                                X  X
                                YPAFK
                                ||||
  GENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRPHLASYPAFKLPDDVNVDPELLQGDDGG
210      220      230      240      250      260  X  270      280

  IAESDGILSLSHPGADRRRAGRYLCRRAEALSY
      290      300      310

```

2. US-08-249-182-3 (1-5)

S29861 prismane protein - *Desulfovibrio vulgaris*

```

ENTRY      S29861      #type complete
TITLE      prismane protein - Desulfovibrio vulgaris
ORGANISM    #formal_name Desulfovibrio vulgaris
DATE        08-Dec-1993; #sequence_revision 08-Dec-1993; #text_change
            08-Dec-1993
ACCESSIONS  S29861
REFERENCE   S29861
#authors    Stokkermans, J.P.W.G.; Pierik, A.J.; Wolbert, R.B.G.; Hagen,
            W.R.; van Dongen, W.M.A.M.; Veeger, C.
#journal     Eur. J. Biochem. (1992) 208:435-442
#title       The primary structure of a protein containing a putative
            [6Fe-6S] prismane cluster from Desulfovibrio vulgaris
            (Hildenborough).
#accession   S29861
##status     preliminary
##residues   1-553 ##label STD
##cross-references EMBL:Z11707
SUMMARY     #length 553 #molecular-weight 60163 #checksum 5194
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.08
Residue Identity   =  100%  Matches          =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

                                X  X
                                YPAFK
                                ||||
  EITQVNIGVGKNPGILISCHDLKDMAELLKQTEGTGVDVYTHGEMLPANYYPAFKKYPHFVGNVGGSWHQQN
      230      240      250      260      270  X  280      290

  PEFESFNGPILLTTNCLVPLKKENTYLDRLYTT
      300      310      320

```

3. US-08-249-182-3 (1-5)

S24389 prismane protein - *Desulfovibrio desulfuricans*

4. A42971	cytochrome P-450terp, P-450te	428	5	5	4.08	0
5. S27653	cytochrome P450-terpredoxin -	428	5	5	4.08	0
6. S37334	hypothetical protein YBL0511	418	5	5	4.08	0
7. S39833	hypothetical protein YBL0511	418	5	5	4.08	0
8. A42329	autotaxin - human (fragments)	114	5	5	4.08	0
9. S12888	DNA-binding protein II - Ther	95	5	5	4.08	0

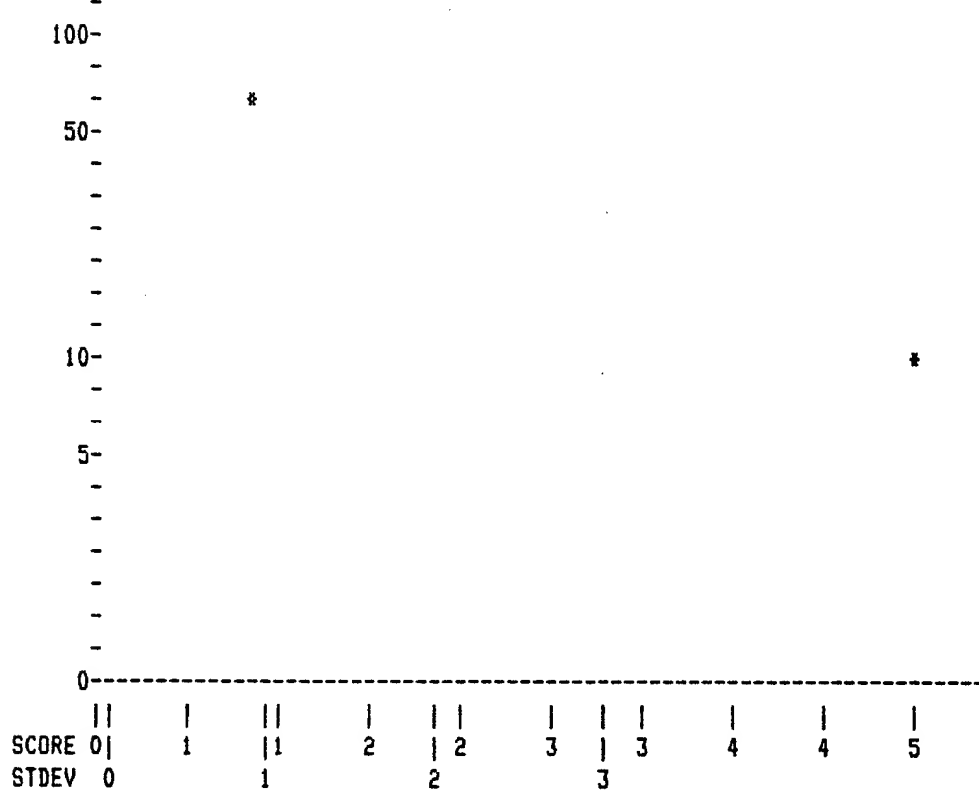
The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
*** 3 standard deviations above mean ***						
10. S35946	araA protein - Escherichia co	43	4	4	3.06	0
11. S10561	photosystem II 21K protein -	66	4	4	3.06	0
12. CCP55S	cytochrome c551 - Pseudomonas	82	4	4	3.06	0
13. CCPB6	cytochrome c6 - Plectonena bo	85	4	4	3.06	0
14. CCA16	cytochrome c6 - Anabaena vari	86	4	4	3.06	0
15. JC1206	DNA-binding protein HU - Baci	90	4	4	3.06	0
16. JC1207	DNA-binding protein HU - Baci	90	4	4	3.06	0
17. DNBS2F	DNA-binding protein HU - Baci	90	4	4	3.06	0
18. S24373	DNA-binding protein HBSu (syn	92	4	4	3.06	0
19. JC1208	DNA-binding protein HU - Baci	92	4	4	3.06	0
20. JC1209	DNA-binding protein HU - Baci	92	4	4	3.06	0
21. S00015	DNA-binding protein HB - Baci	92	4	4	3.06	0
22. S28361	hypothetical protein 12.1 - b	102	4	4	3.06	0
23. S06985	hypothetical protein (nifH1 3	105	4	4	3.06	0
24. 00ECRP	hypothetical protein A-105 -	105	4	4	3.06	0
25. S16667	hypothetical protein P12 - ri	110	4	4	3.06	0
26. B40785	ORF2 protein - rice tungro ba	110	4	4	3.06	0
27. CCA153	cytochrome c6 precursor - Ana	111	4	4	3.06	0
28. S10101	nodulation protein ENOD2 - ga	112	4	4	3.06	0
29. S19492	hypothetical protein YCR078C	114	4	4	3.06	0
30. CCRF2C	cytochrome c2 precursor - Rho	137	4	4	3.06	0
31. A35720	hypothetical 16.1K protein (p	146	4	4	3.06	0
32. S16853	plastoquinol--plastocyanin re	160	4	4	3.06	0
33. JFBYA1	mating hormone alpha-1 - yeas	175	4	4	3.06	0
34. A35309	signal peptidase (EC 3.4.-.-)	179	4	4	3.06	0
35. A34229	signal peptidase (EC 3.4.-.-)	192	4	4	3.06	0
36. WCHCA	coagulogen precursor - Atlant	195	4	4	3.06	0
37. R5SP22	ribosomal protein L22 - spina	199	4	4	3.06	0
38. J01243	coat protein 1 - barley yello	203	4	4	3.06	0
39. A48608	E1 glycoprotein (C-terminal)	207	4	4	3.06	0
40. S24593	hypothetical protein 3 - pota	208	4	4	3.06	0

1. US-08-249-182-3 (1-5)

A26879 alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1

ENTRY A26879 #type complete
TITLE alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
precursor - Klebsiella pneumoniae
ALTERNATE_NAMES pullulanase
ORGANISM #formal_name Klebsiella pneumoniae
DATE 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change
28-Apr-1993
ACCESSIONS A26879
REFERENCE A26879
#authors Katsuragi, N.; Takizawa, N.; Murooka, Y.
#journal J. Bacteriol. (1987) 169:2301-2306
#title Entire nucleotide sequence of the pullulanase gene of
Klebsiella aerogenes W70.
#cross-references MUID:87194626
#contents K. aerogenes, strain W70
#accession A26879



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.98

Times:	CPU	Total Elapsed
	00:01:23.90	00:01:51.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	3917

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

9 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. A26879	alpha-dextrin endo-1,6-alpha-	1096	5	5	4.08	0
2. S29861	prismene protein - Desulfovib	553	5	5	4.08	0

CC native PLRV coat protein; only the DNA sequence.
 CC The modified sequence may be used to transform plants of the Solanaceae
 CC family (e.g. potatoes, tobacco, tomato, pepper) to obtain transgenic
 CC plants resistant to PLRV.
 SQ Sequence 208 AA;
 SQ 10 A; 26 R; 10 N; 7 D; 0 B; 3 C; 12 Q; 6 E; 0 Z; 20 G; 3 H;
 SQ 8 I; 9 L; 12 K; 3 M; 8 F; 11 P; 23 S; 11 T; 2 W; 5 Y; 19 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                YPAFK
                                ||||
RRRRRRGGNRRSRRTGVPRGRGSSETFVFTKDNLVGNSQGSFTFGPSVSDCPAFKDGILKAYHEYKITSILL
  50      60      70      80      90      X 100      110

  QFVSEASSTSPGSIAYELDPHCKVSSLQSYVVK
  120      130      140      150
> 0 <
0| 0 IntelliGenetics
> 0 <

```

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_3p.res made by on Thu 22 Sep 94 10:09:58-PDT.

Query sequence being compared:US-08-249-182-3 (1-5)
 Number of sequences searched: 70848
 Number of scores above cutoff: 3917

Results of the initial comparison of US-08-249-182-3 (1-5) with:
 Data bank : PIR 41, all entries

100000-
 -
 N -
 U50000-
 M -
 B -
 E -
 R *
 -
 D -
 F10000-
 -
 S -
 E 5000-
 Q -
 U -
 E -
 N -
 C -
 E -
 S 1000-
 -
 -
 500-
 -
 -
 -
 -
 -
 -

*

*

*

PT protein - comprises the DNA sequence having at least one internal
PT translation initiation codon in different frame than altered
PT sequence
PS Disclosure; Fig 4; 32pp; English.
CC A synthetic modified PLRV coat protein gene was designed to
CC incorporate nucleotide changes to facilitate the expression of the
CC gene in plants. The modifications make the non-plant gene more
CC "plant-like".
CC Two translation initiation sites (beginning at nucleotides 26 and 32)
CC at the 17 kD ORF in a different reading frame than the native PLRV coat
CC protein are altered (ATG -> ACG). A stronger stop codon (beginning at
CC nucleotide 625) is also provided in the modified PLRV DNA sequence
CC (TAG -> TAA). In addition, a TAG codon is placed behind the TAA
CC termination codon to create a tandem translation stop signal.
CC The modified sequence may be used to transform plants of the Solanaceae
CC family (e.g. potatoes, tobacco, tomato, pepper) to obtain transgenic
CC plants resistant to PLRV.
SQ Sequence 208 AA;
SQ 10 A; 26 R; 10 N; 7 D; 0 B; 3 C; 12 Q; 6 E; 0 Z; 20 G; 3 H;
SQ 8 I; 9 L; 12 K; 3 M; 8 F; 11 P; 23 S; 11 T; 2 W; 5 Y; 19 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
RRRRRRGGNRRSRTGVPRGRGSSETFVFTKDNLVGNSQGSFTFGPSVSDCPAFKDGILKAYHEYKITSILL
50 60 70 80 90 X 100 110
QFVSEASSTSPGSIAYELDPHCKVSSLQSYVNK
120 130 140 150

15. US-08-249-182-3 (1-5)

R32917 Modified PLRV coat protein.

ID R32917 standard; Protein; 208 AA.
AC R32917;
DT 29-JUN-1993 (first entry)
DE Modified PLRV coat protein.
KW Virus; resistance; translation initiation site; stop codon;
KW Solanaceae; transgenic; plant; mutagenesis.
OS Potato leafroll virus isolate Prosser LR7.
PN EP-531273-A.
PD 10-MAR-1993.
PF 02-SEP-1992; 870141.
PR 03-SEP-1991; US-753738.
PA (MONS) MONSANTO CO.
PI Hemenway CL, Lawson EC, Turner NE, Weiss JD;
DR WPI; 93-078751/10.
DR N-PSDB; 037630.
PT Modified DNA sequence encoding potato leaf roll virus coat
PT protein - comprises the DNA sequence having at least one internal
PT translation initiation codon in different frame than altered
PT sequence
PS Disclosure; Fig 3; 32pp; English.
CC The native PLRV coat protein DNA sequence was used to obtain the
CC novel PLRV coat protein DNA sequence with improved virus resistance
CC characteristics. Two translation initiation sites (beginning at
CC nucleotides 26 and 32) at the 17 kD ORF in a different reading frame
CC than the native PLRV coat protein are altered (ATG -> ACG).
CC A stronger stop codon (beginning at nucleotide 625) is also provided
CC in the modified PLRV DNA sequence (TAG -> TAA).
CC The mutagenesis does not alter the amino acid sequence of the

FT /note= "tryptic fragment"
 FT Region 181..184
 FT /note= "tryptic fragment"
 FT Region 189..197
 FT /note= "tryptic fragment"
 FT Region 202..208
 FT /note= "tryptic fragment"
 PN EP-531273-A.
 PD 10-MAR-1993.
 PF 02-SEP-1992; 870141.
 PR 03-SEP-1991; US-753738.
 PA (MONS) MONSANTO CO.
 PI Hemenway CL, Lawson EC, Turner NE, Weiss JD;
 DR WPI; 93-078751/10.
 DR N-PSDB; Q37629.
 PT Modified DNA sequence encoding potato leaf roll virus coat
 PT protein - comprises the DNA sequence having at least one internal
 PT translation initiation codon in different frame than altered
 PT sequence
 PS Disclosure; Fig 1; 32pp; English.
 CC The native PLRV coat protein DNA sequence was used to obtain the
 CC novel PLRV coat protein DNA sequence with improved virus resistance
 CC characteristics. Two translation initiation sites (beginning at
 CC nucleotides 26 and 32) at the 17 kD ORF in a different reading frame
 CC than the native PLRV coat protein are pref. altered. A stronger stop
 CC codon is also provided in the modified PLRV DNA sequence, i.e. TAA TAG.
 CC The modified sequence may be used to transform plants of the Solanaceae
 CC family (e.g. potatoes, tobacco, tomato, pepper) to obtain transgenic
 CC plants resistant to PLRV.
 SQ Sequence 208 AA;
 SQ 10 A; 26 R; 10 N; 7 D; 0 B; 3 C; 12 Q; 6 E; 0 Z; 20 G; 3 H;
 SQ 8 I; 9 L; 12 K; 3 M; 8 F; 11 P; 23 S; 11 T; 2 W; 5 Y; 19 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 RRRRRRCGNRRSRRTGVPRGRGSSETFVFTKDNLVGNSQGSFTFGPSVSDCPAFKDGILKAYHEYKITSILL
 50 60 70 80 90 X 100 110
 QFVSEASSTSPGSIAYELDPHCKVSSLQSYVVK
 120 130 140 150

14. US-08-249-182-3 (1-5)

R32918 Synthetic modified PLRV coat protein.

ID R32918 standard; Protein; 208 AA.
 AC R32918;
 DT 29-JUN-1993 (first entry)
 DE Synthetic modified PLRV coat protein.
 KW Virus; resistance; translation initiation site; stop codon;
 KW Solanaceae; transgenic; plant.
 OS Synthetic.
 PN EP-531273-A.
 PD 10-MAR-1993.
 PF 02-SEP-1992; 870141.
 PR 03-SEP-1991; US-753738.
 PA (MONS) MONSANTO CO.
 PI Hemenway CL, Lawson EC, Turner NE, Weiss JD;
 DR WPI; 93-078751/10.
 DR N-PSDB; Q37631.
 PT Modified DNA sequence encoding potato leaf roll virus coat

12. US-08-249-182-3 (1-5)

R44505 PLRV integument protein.

ID R44505 standard; Protein; 208 AA.
 AC R44505;
 DT 21-JUN-1994 (first entry)
 DE PLRV integument protein.
 KW Primer; isolation; potato leaf roll virus; PLRV; integument protein;
 KW transformation; plasmid; potato; electroporation; agrobacterium;
 KW microinjection; redifferentiation; resistance.
 OS Potato leaf roll virus.
 PN J05304847-A.
 PD 19-NOV-1993.
 PF 30-APR-1992; 135561.
 PR 30-APR-1992; JP-135561.
 PA (HOKK-) HOKKAIDO GURIINBAIO KENKYUSHO KK.
 DR WPI; 93-408237/51.
 DR N-PSDB; 053460.
 PT Potatoes resistant against potato leaf roll virus - comprises
 PT transforming potato cells using recombinant vector combined DNA
 PS Disclosure; Page 7; 8pp; Japanese.
 CC This sequence is encoded by the potato leaf roll virus (PLRV)
 CC integument protein gene. The PLRV integument protein gene may be
 CC transformed into a microorganism which is then cultured. The cultured
 CC plasmid may be introduced in to a potato cell by electroporation,
 CC agrobacterium transformation or by microinjection. Expression of the
 CC integument protein causes redifferentiation of transformed cells. This
 CC is used in the production of potatoes resistant to PLRV.
 SQ Sequence 208 AA;
 SQ 10 A; 26 R; 10 N; 7 D; 0 B; 3 C; 12 Q; 6 E; 0 Z; 20 G; 3 H;
 SQ 8 I; 10 L; 12 K; 3 M; 8 F; 9 P; 24 S; 12 T; 2 W; 5 Y; 18 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 RRRRRRGGNRRSRRGTGVPGRGSSSETFVFTKDNLVGNSQGSFTFGPSLSDCPAFKDGILKAYHEYKITSILL
 50 60 70 80 90 X 100 110
 QFVSEASSTSSGSIAYELDPHCKVSSLQSYVNK
 120 130 140 150

13. US-08-249-182-3 (1-5)

R32916 Native PLRV coat protein.

ID R32916 standard; Protein; 208 AA.
 AC R32916;
 DT 29-JUN-1993 (first entry)
 DE Native PLRV coat protein.
 KW Virus; resistance; translation initiation site; stop codon;
 KW Solanaceae; transgenic; plant.
 OS Potato leafroll virus isolate Prosser LR7.
 FH Key Location/Qualifiers
 FT Region 32..42
 FT /note= "tryptic fragment"
 FT Region 77..91
 FT /note= "tryptic fragment"
 FT Region 141..151

02-NOV-1988; 003877.
PR 04-NOV-1987; US-117099.
PA (CALB-) California Biotechn.
PI Benson BJ, White RT, Schilling JW, Buckley D, Scarborough RM,
DR WPI; 89-165617/22.
DR N-PSDB; N90107.
PT Human SP-18 and SP-5 derived peptide(s) -
PT with alveolar surfactant protein activity, used for treating
PT respiratory distress syndrome, pneumonia and bronchitis
PS Disclosure; Fig 9; .pp; English.
CC ASP proteins including the hSP-18- and hSP-5-derived peptides can be
CC used as a carrier or vehicle for delivery of other active and important
CC molecules to and/or through the lung to the blood vasculature.
SQ Sequence 187 AA;
SQ 11 A; 5 R; 7 N; 7 D; 0 B; 5 C; 8 Q; 11 E; 0 Z; 7 G; 11 H;
SQ 12 I; 14 L; 11 K; 6 M; 15 F; 7 P; 8 S; 10 T; 3 W; 8 Y; 21 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
TGYYTVDISQWHRKEHFEAFQSVAGCTYNQTVQLDITAFLLKTVKKNKHKFYPAFIHILARLMNAHPEFRMAM
10 20 30 40 50 X 60 70

KDGELVINDSVHPCYTVFHEQTETFSLLWSEYH
80 90 100 110

11. US-08-249-182-3 (1-5)
P90731 Limulus polyphemus coagulogen.

ID P90731 standard; protein; 195 AA.
AC P90731;
DT 13-OCT-1989 (first entry)
DE Limulus polyphemus coagulogen.
KW Coagulogen; Limulus polyphemus; haematocytes; coagulation.
OS Limulus polyphemus.
PN J01160484-A.
PD 23-JUN-1989.
PF 17-DEC-1987; 317425.
PR 17-DEC-1987; JP-317425.
PA (GREC) Green Cross Corp.
DR WPI; 89-224282/31.
PT Coagulogen cDNA
PT - which originates from Limulus polyphemus and is present mainly
PT in haematocytes.
PS Claim 1; fig 1; 15pp; Japanese.
CC Coagulogen from Limulus polyphemus (see corresp. N90363). Is present
CC mainly in haematocytes, and is involved in body humor coagulation.
SQ Sequence 195 AA;
SQ 9 A; 13 R; 5 N; 6 D; 0 B; 16 C; 8 Q; 15 E; 0 Z; 14 G; 4 H;
SQ 9 I; 11 L; 11 K; 1 M; 13 F; 10 P; 13 S; 14 T; 1 W; 5 Y; 17 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
NVPTCLCEPTLLGRKVIVSQETKDKIEEAVQAITDKDEISGRGFSIFGGHPAFKECGKYECRTVTSEDSRC
30 40 50 60 70 X 80 90

ID R15612 standard; Protein; 187 AA.
AC R15612;
DT 16-MAR-1992 (first entry)
DE SP-C from pC149SP-C insert.
KW Alveolar surfactant protein; CAT; fusion protein; SP-5.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Protein 1..149
FT /label= CAT
FT Region 150..152
FT /label= linker
FT Protein 152..187
FT /label= SP-C
PN W09118015-A.
PD 28-NOV-1991.
PF 17-MAY-1991; U03490.
PR 17-MAY-1990; US-524360.
PA (CALB-) CALIF BIOTECHN INC.
PI Benson BJ, White RT, Schilling JW, Buckley DI, Scarborough RM;
DR WPI; 91-369185/50.
DR N-PSDB; Q15266.
PT New alveolar surfactant protein analogues - used for treating
PT respiratory distress syndrome, pneumonia and bronchitis
PS Disclosure; Fig 9; 67pp; English.
CC An amino acid sequence of a 187 residue fusion protein encoded
CC by pC149SP-C is a slight modification of the sequence shown in
CC Q15265. In plasmid pC149SP-C, the 149 amino acids of CAT are
CC joined to 35 amino acids of SP-5 through a linker of 3 amino acids.
CC The SP-5 comprises 18.7% of the total fusion.
CC To construct pC149SP-C, a portion of the CAT segment of pC210SP-C
CC extending from the DdeI site at nucleotide 523 (Q15265) to the
CC EcoRI site at nucleotide 728 was removed and replaced by a set of
CC two complementary oligonucleotides.
CC See also R15602-7, Q15262-63 and Q15265-66.
SQ Sequence 187 AA;
SQ 11 A; 5 R; 7 N; 7 D; 0 B; 5 C; 8 Q; 11 E; 0 Z; 7 G; 11 H;
SQ 12 I; 15 L; 11 K; 6 M; 15 F; 7 P; 8 S; 10 T; 3 W; 8 Y; 20 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X

X

YPAFK

||||

TGYTTVDISQWHRKEHFEAFQSVAGCTYNQTVQLDITAF

1020304050X6070

LKTVKKNKHKFYPAFIHILARLMNAHPEFRMAM

KDGELVIWDSVHPCYTVFHEQTETFFSSLWSEYH

8090100110

10. US-08-249-182-3 (1-5)
P98407 Sequence of surfactant protein CAT-SP-5 fusion pro

ID P98407 standard; protein; 187 AA.
AC P98407;
DT 19-JAN-1991 (first entry)
DE Sequence of surfactant protein CAT-SP-5 fusion protein encoded by
DE pC149SP-C
KW Alveolar surfactant protein; respiratory distress syndrome;
KW pneumonia; bronchitis; therapy.
OS Homo sapiens.
PN W08904326-A.
PD 18-MAY-1989.

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                X  X
                YPAFK
                || ||
QLTKPRPVILDPADPTGNLGGDPKGWRQLAGEAEAWLNYPCKNWDGSPVSSWILLAESNSTDDDEDDPRT
   10      20      30      40  X  50      60      70

YQKYGYIGTHEYPHFSHRPSTL
   80      90

```

8. US-08-249-182-3 (1-5)

R05445 CAT-A4-751i hybrid protein.

ID R05445 standard; protein; 132 AA.
 AC R05445;
 DT 30-JUL-1990 (first entry)
 DE CAT-A4-751i hybrid protein.
 KW CAT; hybrid protein; A4-751i protein.
 OS Synthetic
 FH Key Location/Qualifiers
 FT Peptide 1..76
 FT /label=CAT
 FT Peptide 77..132
 FT /label=M-751i
 PN W09001540-A.
 PD 22-FEB-1990.
 PF 09-AUG-1989; U03417.
 PR 11-AUG-1988; US-231224.
 PA (CALB-) Calif Biotechn Inc.
 PI Hilliker S. White R;
 DR WPI; 90-083499/11.
 DR N-PSDB; 003563.
 PT Heterologous protein expression on prokaryotic host -
 PT using 3' truncated chloramphenicol acetyl transferase gene to
 PT stably express hybrid protein.
 PS Example; Fig 4A; 67pp; English.
 CC The N-terminus of the CAT protein joined to the synthetic A4-751i gene
 CC product preceded by a chemical cleavage site (hydroxylamine)
 CC encoded by Asn-Gly.
 CC See also 003557 to 005366; and 004767.
 SQ Sequence 132 AA;
 SQ 11 A; 5 R; 7 N; 4 D; 0 B; 7 C; 6 Q; 10 E; 0 Z; 9 G; 5 H;
 SQ 7 I; 4 L; 9 K; 4 M; 10 F; 4 P; 5 S; 10 T; 2 W; 6 Y; 7 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

                X  X
                YPAFK
                ||||
TGYYTVDISQWHRKEHFEAFQSVAGCTYNQTVQLDITAFKTVKKNKHKFYPAFIHILARLMNAHPEFNGEV
   10      20      30      40      50  X  60      70

CSEQAETGPCRAMISRWFYFDVTEGKCAPFFYGG
   80      90     100     110

```

9. US-08-249-182-3 (1-5)

R15612 SP-C from pC149SP-C insert.

PR 08-APR-1988; IL-078445.
 PA (YEDA) YEDA RES & DEV CD LTD.
 PI Revel M, Chebath J;
 DR WPI; 87-095196/14.
 DR N-PSDB; N70146.
 PT Recombinant enzyme having (2'-5') oligo A synthetase activity -
 PT used for monitoring the response of a patient to an interferon
 PS Example; Fig 1B; 90pp; English.
 CC A partial cDNA clone (E1) for the DAS mRNA from human SV80 cells was
 CC first obtd. through its ability to select by hybridisation a mRNA
 CC producing OAS activity upon translation in Xenopus laevis oocytes.
 CC The E1 cDNA insert hybridises to 3RNA species of 1.6, 1.8 and 3.6 kb
 CC which are coinduced by IFN in SV80 cells. cDNA clones for the 1.6
 CC and 1.8 RNAs have been isolated and sequenced.
 SQ Sequence 110 AA;
 SQ 7 A; 6 R; 5 N; 5 D; 0 B; 2 C; 5 Q; 5 E; 0 Z; 6 G; 1 H;
 SQ 7 I; 13 L; 6 K; 0 M; 4 F; 13 P; 5 S; 4 T; 5 W; 6 Y; 5 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 || ||
 KNPIIEKYLRRLTKPRPVILDPADPTGNLGGDPKGWRQLAQEAELNYPCKNWDGSPVSSWILLVRPP
 30 40 50 60 70 X 80 90
 ASSLPFIPAPLHEA
 100 110

7. US-08-249-182-3 (1-5)

P71705 Partial (2'-5') oligo A synthetase sequence.

ID P71705 standard; Protein; 111 AA.
 AC P71705;
 DT 18-APR-1991 (first entry)
 DE Partial (2'-5') oligo A synthetase sequence.
 KW In vivo interferon assay; OAS.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Modified -site 61
 FT /label= glycosylated Asn
 PN EP-217102-A.
 PD 08-APR-1987.
 PF 21-AUG-1986; 111585.
 PR 28-AUG-1985; IL-076233.
 PR 08-APR-1986; IL-078445.
 PA (YEDA) YEDA RES & DEV CD LTD.
 PI Revel M, Chebath J;
 DR WPI; 87-095196/14.
 DR N-PSDB; N71375.
 PT Recombinant enzyme having (2'-5') oligo A synthetase activity -
 PT used for monitoring the response of a patient to an interferon
 PS Claim 3; Fig 7B; 90pp; English.
 CC A partial cDNA clone (E1) for the DAS mRNA from human SV80 cells was
 CC first obtd. through its ability to select by hybridisation a mRNA
 CC producing OAS activity upon translation in Xenopus laevis oocytes.
 CC The E1 cDNA insert hybridises to 3RNA species of 1.6, 1.8 and 3.6 kb
 CC which are coinduced by IFN in SV80 cells. cDNA clones for the 1.6
 CC and 1.8 RNAs have been isolated and sequenced. This amino acid
 CC sequence is deduced from the 1.8E clone which comprises the last 5
 CC codons of exon 6 and all of exons 7a and 8.
 SQ Sequence 111 AA;
 SQ 8 A; 4 R; 4 N; 9 D; 0 B; 2 C; 6 Q; 7 E; 0 Z; 8 G; 3 H;

5. US-08-249-182-3 (1-5)

R05446 CAT-GLP-1 hybrid protein.

ID R05446 standard; protein; 104 AA.
 AC R05446;
 DT 30-JUL-1990 (first entry)
 DE CAT-GLP-1 hybrid protein.
 KW CAT; hybrid protein; GLP-1 protein.
 OS Synthetic
 FH Key Location/Qualifiers
 FT Peptide 1..64
 FT /label=CAT
 FT Peptide 65..105
 FT /label=GLP-1 (7-37)
 PN W09001540-A.
 PD 22-FEB-1990.
 PF 09-AUG-1989; U03417.
 PR 11-AUG-1988; US-231224.
 PA (CALB-) Calif Biotechn Inc.
 PI Hilliker S, White R;
 DR WPI; 90-083499/11.
 DR N-PSDB; 004767.
 PT Heterologous protein expression on prokaryotic host -
 PT using 3' truncated chloramphenicol acetyl transferase gene to
 PT stably express hybrid protein.
 PS Example; Fig 4B; 67pp; English.
 CC First 73 amino acids of the CAT protein joined in-frame to
 CC the synthetic GLP-1 gene product preceded by a Met codon.
 CC (Cyanogen bromide cleavage site.)
 CC See also 003557 to 005366.
 SQ Sequence 104 AA;
 SQ 10 A; 3 R; 3 N; 3 D; 0 B; 1 C; 6 Q; 7 E; 0 Z; 4 G; 6 H;
 SQ 6 I; 6 L; 10 K; 3 M; 8 F; 2 P; 5 S; 9 T; 2 W; 4 Y; 6 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||

TCYTTVDISQWHRKEHFEAFQSVAGCTYNGTVQLDITAFLLKTVKKNKHKFYPAFIHILARLMNAHPEFMHAE

10 20 30 40 50 X 60 70

GTFTSDVSSYLEGQAAKEFIAWLKGR

80 90 100

6. US-08-249-182-3 (1-5)

P70092 Sequence encoded by (2'-5') oligo A synthetase (OA

ID P70092 standard; Protein; 110 AA.
 AC P70092;
 DT 18-APR-1991 (first entry)
 DE Sequence encoded by (2'-5') oligo A synthetase (OAS) E1 cDNA
 DE clone 174-3.
 KW In vivo interferon assay.
 OS Homo sapiens.
 PN EP-217102-A.
 PD 08-APR-1987.
 PF 21-AUG-1986; 111585.
 PR 28-AUG-1985; IL-076233.

such control of the crystal growth providing a degree of plant frost tolerance. The peptides have mol.wts. of 5 kD, 9 kD, 11 kD, 22 kD, 24 kD, 30 kD, 36 kD, 60 kD and 68 kD. The N-terminals of the 30 kD, 11 kD and 9 kD peptides are given in R31179-81 respectively.

Sequence 20 AA;
3 A; 0 R; 1 N; 0 D; 0 B; 1 C; 1 Q; 0 E; 0 Z; 3 G; 0 H;
2 I; 1 L; 0 K; 0 M; 2 F; 4 P; 0 S; 0 T; 0 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||
AIFCGQVNPALGPPIYPAFG
10 X 20

4. US-08-249-182-3 (1-5)

R32366 CAT 1-73 peptide.

ID R32366 standard; Protein; 73 AA.
AC R32366;
DT 18-JUN-1993 (first entry)
DE CAT 1-73 peptide.
KW Human; proinsulin; vector; pUC19; pPINS; CAT; pUC-CAT-proinsulin;
KW insulin analogue; type I; type II; diabetes.
OS Synthetic.
PN W09303174-A.
PD 18-FEB-1993.
PF 31-JUL-1992; U06451.
PR 08-AUG-1991; US-741938.
PR 30-JUL-1992; US-918953.
PA (PFIZ) PFIZER INC.
PA (SCIO-) SCIOS INC.
PI Andy RJ, Larson ER;
DR WPI; 93-076530/09.
DR N-PSDB; Q37002.
PT New hepato selective and peripheral selective human insulin
PT analogues - and their corresp. DNA, for treatment of type I and
PT type II diabetes
PS Disclosure; Fig 2a; 58pp; English.
CC The sequence given represents amino acids 1-73 of CAT. This sequence
CC was used in the construction of proinsulin analogues. The sequences
CC in Q36996-7001 are oligonucleotides which can be combined to form a
CC gene which codes for human proinsulin. The resulting cDNA coding for
CC proinsulin was inserted into plasmid vector pUC19 and digested with
CC KpnI and HindIII. This resulted in the formation of the vector pPINS.
CC This CAT fragment was inserted into pPINS to give a plasmid which
CC contained DNA sequences which coded for amino acids 1-73 of CAT, an 8
CC amino acid linker sequence and human proinsulin (see Q37003). This
CC plasmid, pUC-CAT-proinsulin, could be used in the formation of insulin
CC analogues which may be used in the treatment of types I and II
CC diabetes.
SQ Sequence 73 AA;
SQ 6 A; 2 R; 3 N; 2 D; 0 B; 1 C; 5 Q; 4 E; 0 Z; 1 G; 5 H;
SQ 5 I; 4 L; 8 K; 2 M; 6 F; 2 P; 2 S; 7 T; 1 W; 3 Y; 4 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.86
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
||||

2. US-08-249-182-3 (1-5)
P82507 Pullulanase protein.

ID P82507 standard; protein; 1096 AA.
AC P82507;
DT 01-NOV-1990 (first entry)
DE Pullulanase protein.
KW Pullulanase; starch; alcohol prodn.
OS Klebsiella aerogenes.
PN J63245676-A.
PD 12-OCT-1988.
PF 31-MAR-1987; 078355.
PR 31-MAR-1987; JP-078355.
PA (ELED) Denki Kagaku Kogyo KK; (SUNR) Suntory Ltd.
DR WPI; 88-333488/47.
DR N-PSDB; N81341.
PT Gene encoding pullulanase - derived from recombinant plasmid pMP1 contg.
PT gene from Klebsiella genus.
PS Disclosure; p; Japanese.
CC The pullulanase protein cleaves alpha-1,6-glucoside bonds of starch and
CC is effective in decomposition of branched starch. It is used in the
CC prodn. of maltose and glucose from starch, and of alcohol from starch
CC via glucose. Amino acid residues 1-19 can be deleted.
SQ Sequence 1096 AA;
SQ 108 A; 63 R; 42 N; 93 D; 0 B; 10 C; 51 Q; 42 E; 0 Z; 89 G; 22 H;
SQ 38 I; 86 L; 42 K; 21 M; 31 F; 52 P; 96 S; 70 T; 18 W; 41 Y; 81 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.82
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||||

GENKPIVRLYVSHSSKVAADSNGEFS DKYVKLTPTTVNQGVSMRFPHLASYPAFKLPDDVNVDELLOGDDGG
210 220 230 240 250 260 X 270 280

IAESDGI LSLSHPGADRRRAGRYLCRRAEALSY
290 300 310

3. US-08-249-182-3 (1-5)
R31181 N-terminal of 9 kD antifreeze polypeptide.

ID R31181 standard; peptide; 20 AA.
AC R31181;
DT 14-MAY-1993 (first entry)
DE N-terminal of 9 kD antifreeze polypeptide.
KW Frost tolerance; plant; ice crystal; cryopreservation; winter rye.
OS Secale cereale cv. Musketeer.
PN W09222581-A.
PD 23-DEC-1992.
PF 12-JUN-1992; CA0255.
PR 13-JUN-1991; GB-012774.
PR 13-DEC-1991; GB-026485.
PA (UYWA-) UNIV WATERLOO.
PI Griffith M;
DR WPI; 93-018083/02.
PT Polypeptide(s) responsible for frost tolerance in plants - used
PT in cryo preservation of biological tissues and for improving
PT quality of frozen foods
PS Disclosure; Page 25; 52pp; English.
CC Antifreeze polypeptides common to frost tolerant plants were
CC isolated. The peptides are located in extracellular spaces of plant
CC cells to control ice crystal growth in the intercellular plant space.

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.04

Times:	CPU	Total Elapsed
	00:00:24.91	00:00:25.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	4652

Cut-off raised to 2.
Cut-off raised to 3.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37445	Autotaxin peptide ATX 20.	5	5	5	4.82	0

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
2. PB2507	Pullulanase protein.	1096	5	5	4.82	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 3 standard deviations above mean ****						
3. R31181	N-terminal of 9 kD antifreeze	20	4	4	3.86	0
4. R32366	CAT 1-73 peptide.	73	4	4	3.86	0
5. R05446	CAT-GLP-1 hybrid protein.	104	4	4	3.86	0
6. P70092	Sequence encoded by (2'-5') o	110	4	4	3.86	0
7. P71705	Partial (2'-5') oligo A synth	111	4	4	3.86	0
8. R05445	CAT-A4-751i hybrid protein.	132	4	4	3.86	0
9. R15612	SP-C from pC149SP-C insert.	187	4	4	3.86	0
10. P98407	Sequence of surfactant protei	187	4	4	3.86	0
11. P90731	Limulus polyphenus coagulogen	195	4	4	3.86	0
12. R44505	PLRV integument protein.	208	4	4	3.86	0
13. R32916	Native PLRV coat protein.	208	4	4	3.86	0
14. R32918	Synthetic modified PLRV coat	208	4	4	3.86	0
15. R32917	Modified PLRV coat protein.	208	4	4	3.86	0
16. R13586	PLRV capsid protein.	208	4	4	3.86	0
17. R05101	Potato leaf roll virus 23K co	208	4	4	3.86	0
18. R05425	Amino acid sequence for a CA	240	4	4	3.86	0
19. P92070	Sequence of chloramphenicol a	241	4	4	3.86	0

X X
X
YPAFK
||||
YPAFK
X X

Initial Score = 5 Optimized Score = 5 Significance = 4.82
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

SA 0 I: 0 L: 1 K: 0 M: 1 F: 1 P: 0 S: 0 T: 0 W: 1 Y: 0 V:
SA 1 A: 0 R: 0 N: 0 D: 0 B: 0 C: 0 Q: 0 E: 0 Z: 0 G: 0 H:
Sequence 5 AA:
CC therapy.
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC fluids can be used to predict disease outcomes and/or choice of
CC presence of autotaxin. The level of autotaxin in tissue or body
CC metastasis and in immunostains of patient samples to detect the
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC The sequence is that of autotaxin peptide ATX 20. It may be used to
PS Example: Page 33; 36pp; English.
PT diagnosis and therapy
PT Motility stimulating protein named autotaxin - useful in cancer
DR WPI: 93-085861/10.
PI Kruttsch H, Liotta LA, Schiffmann E, Stracke M.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PR 17-JAN-1992; US-822043.
PF 17-JAN-1992; 822043.
PD 01-JAN-1993.
PN US7822043-A.
OS Synthetic.
KM cancer therapy; crosslinked toxins.
KM immunostains; disease outcome prediction; therapy choice;
KM Cell motility stimulating; cancer metastasis; antibody; detection;
DE Autotaxin peptide ATX 20.
DT 22-JUL-1993 (first entry)
AC R37445;
ID R37445 standard; peptide: 5 AA.

1. US-08-249-182-3 (1-5)
R37445 Autotaxin peptide ATX 20.

40. R01940	Tumour necrosis factor.	417	4	4	3.86	0
39. P90119	Heat stable sarcosine oxidase	387	4	4	3.86	0
38. R22623	Thermus aquaticus xylose isom	387	4	4	3.86	0
37. R38078	Sarcosine oxidase M.	387	4	4	3.86	0
36. P70457	Sequence of gpc encoded by se	377	4	4	3.86	0
35. R15272	Fusarium oxysporum DSM 2672 e	376	4	4	3.86	0
34. R25527	Fusarium oxysporum DSM 2672 e	376	4	4	3.86	0
33. R25466	Endoglucanase #2.	376	4	4	3.86	0
32. R25429	Cellulase contained in a dete	376	4	4	3.86	0
31. R27969	Endoglucanase enzyme.	376	4	4	3.86	0
30. R37151	Dye transfer inhibiting comps	376	4	4	3.86	0
29. R42064	Endoglucanase enzyme.	376	4	4	3.86	0
28. P70094	Sequence encoded by (2'-5') o	364	4	4	3.86	0
27. R34102	Bacterial delta-b-desaturase.	359	4	4	3.86	0
26. R14408	Nuclear factor C/EBP2.	345	4	4	3.86	0
25. R05418	CAT:SP-B hybrid protein.	293	4	4	3.86	0
24. R22903	S.cremoris Abi 105 phage resi	274	4	4	3.86	0
23. R05419	CAT:SP-C hybrid protein.	251	4	4	3.86	0
22. R15611	SP-C from pC210SP-C insert.	251	4	4	3.86	0
21. P92068	Fusion protein comprising chl	250	4	4	3.86	0
20. P11218	nonan VI of mouse laminin RI	248	4	4	3.86	0

RA HAY J.M.; JONES M.C.; BLAKEBROUGH M.; DASGUPTA I.; DAVIES J.W.;
 RA HULL R.;
 RL NUCLEIC ACIDS RES. 19:2615-2621(1991).
 DR EMBL; X57924; RTBVPHIL.
 DR PIR; S16667; S16667.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 110 AA; 11910 MW; 64722 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      YPAFK
      || ||
MSADYPTFKALEKFKNLES DTAGKDKFNWVFTLENIKSAADVNLASKGLVQLYALQEI
      X  10      20      30      40      50
  
```

15. US-08-249-182-3 (1-5)
 YP12_RTBV HYPOTHETICAL P12 PROTEIN (ORF 2).

ID YP12_RTBV STANDARD; PRT; 110 AA.
 AC P27529;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL P12 PROTEIN (ORF 2).
 OS RICE TUNGRO BACILLIFORM VIRUS (RTBV).
 OC VIRIDAE; NOT YET CLASSIFIED.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92024093
 RA QU R.D., BHATTACHARYYA M., LACO G.S., DE KOCHKO A., RAD B.L.,
 RA KANIEWSKA M.B., ELMER J.S., ROCHESTER D.E., SMITH C.E.,
 RA BEACHY R.N.;
 RL VIROLOGY 185:354-364(1991).
 DR EMBL; M65026; LERTUORFS.
 DR PIR; B40785; B40785.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 110 AA; 11924 MW; 64761 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X  X
      YPAFK
      || ||
MSADYPTFKALEKFKNLES DTAGKDKFNWVFTLENIKTAADVNLASKGLVQLYALQEI
      X  10      20      30      40      50
  
```

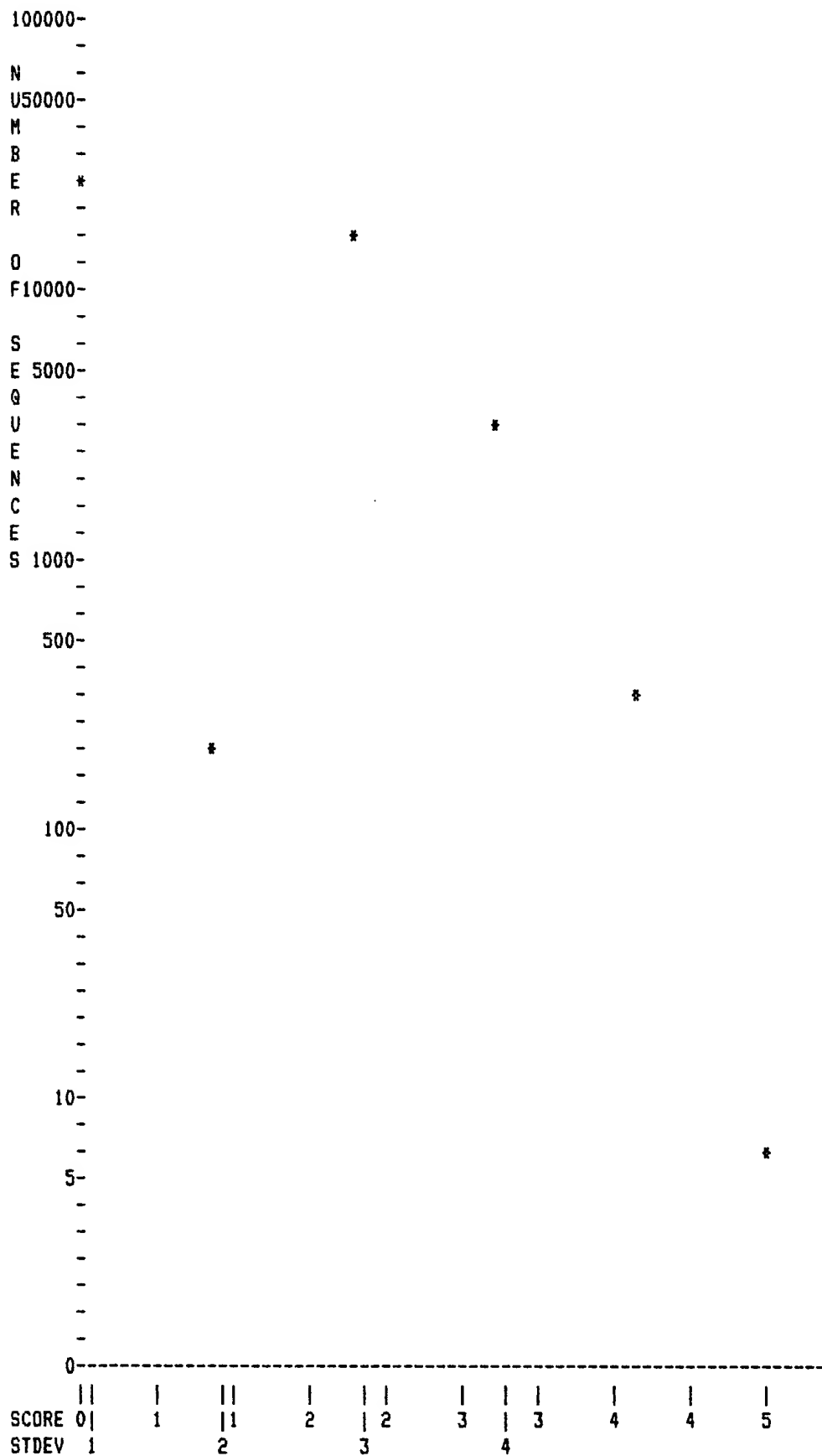
> 0 <
 0| 10 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4 *Seq. 4*

Results file u249_4a.res made by on Thu 22 Sep 94 10:12:35-PDT.

Query sequence being compared:US-08-249-182-4 (1-5)
 Number of sequences searched: 42145
 Number of scores above cutoff: 3812

Results of the initial comparison of US-08-249-182-4 (1-5) with:



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.12

Times:	CPU	Total Elapsed
	00:00:26.98	00:00:29.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	3812

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. R37446	Autotaxin peptide ATX 24.	5	5	5	4.48	0

5 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
2. R42848	VIP receptor protein.	459	5	5	4.48	0
3. R26150	A5B7 gH-2 antibody grafted he	146	5	5	4.48	0
4. R20793	CDR-grafted, humanised heavy	146	5	5	4.48	0
5. R41469	MAB 25D2 humanised heavy chai	140	5	5	4.48	0
6. R40179	Humanised antibody CMX5-3 hea	135	5	5	4.48	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 3 standard deviations above mean ****						
7. R41113	HCV peptide XIV or HCV8 (aa 1	22	4	4	3.58	0
8. R41186	HCV NS4 protein HCV7/8.	26	4	4	3.58	0
9. P90810	Synthetic HTLV-1 peptide anti	26	4	4	3.58	0
10. R40053	Hib OMP P1 peptide HIBP1-1 (1	30	4	4	3.58	0
11. R13186	Peptide VII immunoreactive wi	38	4	4	3.58	0
12. R06408	HTLV-1 corresponding peptide	38	4	4	3.58	0
13. R20772	Peptide V based on immunoreac	40	4	4	3.58	0
14. R13185	Peptide (V) immunoreactive wi	40	4	4	3.58	0
15. R13591	HTLV-1 env precursor epitope	40	4	4	3.58	0
16. R33871	Polypeptide p1684 comprising	67	4	4	3.58	0
17. R13345	P1684 HCV antigen (1684-1750)	67	4	4	3.58	0
18. R10026	ompA Signal peptide used to c	78	4	4	3.58	0
19. R13557	HCV C-100 protein immunodonin	90	4	4	3.58	0
20. R13962	Putidaredoxin.	107	4	4	3.58	0
21. R33872	Polypeptide p1689 comprising	117	4	4	3.58	0
22. P92019	Sequence of the polypeptide e	117	4	4	3.58	0
23. P90136	Sequence of hepatitis C virus	117	4	4	3.58	0
24. R13354	P1689 HCV antigen (1689-1805)	118	4	4	3.58	0
25. P92018	Sequence of the polypeptide e	128	4	4	3.58	0

26. P90135	Sequence of hepatitis C virus	128	4	4	3.58	0
27. R25113	Non-A, Non-B Hepatitis Virus	134	4	4	3.58	0
28. R38093	nodC N-terminal portion.	153	4	4	3.58	0
29. P93319	Amino acid sequence of swine	193	4	4	3.58	0
30. P93318	Amino acid sequence of swine	193	4	4	3.58	0
31. R10650	Adenylate kinase.	194	4	4	3.58	0
32. P90745	Recombinant human cardiac myo	195	4	4	3.58	0
33. P91391	Human ventricular myosin ligh	195	4	4	3.58	0
34. R43885	Consensus sequence of C-termi	199	4	4	3.58	0
35. R43879	C-terminal portion of FIPV sp	200	4	4	3.58	0
36. R13897	Nitrile hydratase alpha subun	207	4	4	3.58	0
37. R11717	ENV93/HTLV-1-II fusion protei	217	4	4	3.58	0
38. R33714	A29379.	219	4	4	3.58	0
39. R10102	gp46/p21E env fusion protein	219	4	4	3.58	0
40. P71662	Sequence encoded by adult T-c	228	4	4	3.58	0

1. US-08-249-182-4 (1-5)

R37446 Autotaxin peptide ATX 24.

ID R37446 standard; peptide; 5 AA.
AC R37446;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 24.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 24. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 5 AA;
SQ 1 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 1 S; 0 T; 0 W; 0 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

X  X
QAEVS
|||||
QAEVS
X  X

```

2. US-08-249-182-4 (1-5)

R42848 VIP receptor protein.

ID R42848 standard; Protein; 459 AA.
AC R42848;

DT 13-MAY-1994 (first entry)
 DE VIP receptor protein.
 KW vasoactive intestinal polypeptide receptor; VIP; rat; binding;
 KW adenylate cyclase activity; stimulus.
 OS Rattus rattus.
 PN J05255394-A.
 PD 05-OCT-1993.
 PF 13-FEB-1992; 026607.
 PR 13-FEB-1992; JP-026607.
 PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.
 DR WPI; 93-348480/44.
 DR N-PSDB; 050349.
 PT Vasoactive intestinal polypeptide - prepd. in large amt. by
 PT culturing microbe transformed by new DNA coding polypeptide
 PS Claim 2; Page 6; 14pp; Japanese.
 CC The sequence can be used to produce large amounts of the VIP
 CC receptor peptide, by culturing a microorganism transformed by
 CC the sequence.
 SQ Sequence 459 AA;
 SQ 32 A; 22 R; 14 N; 11 D; 0 B; 17 C; 22 Q; 20 E; 0 Z; 27 G; 11 H;
 SQ 31 I; 52 L; 15 K; 10 M; 30 F; 21 P; 40 S; 20 T; 15 W; 16 Y; 33 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 RKWRRWHLQGVLCWSSKSGHPWGGSNATCSTQVSMLTRVSPSARRSSSFQAEVSLV
 410 420 430 440 450 X X

3. US-08-249-182-4 (1-5)

R26150 A5B7 gH-2 antibody grafted heavy chain.

ID R26150 standard; Protein; 146 AA.
 AC R26150;
 DT 03-FEB-1993 (first entry)
 DE A5B7 gH-2 antibody grafted heavy chain.
 KW humanised antibody; chimaeric; carcino-embryonic antigen; therapy;
 KW diagnosis; carcinomas; CDR; IgG; human; murine; ss.
 OS Chimaeric.
 FH Key Location/Qualifiers
 FT Region 26..35
 FT /note= "grafted murine CDR1"
 FT Region 50..65
 FT /note= "grafted murine CDR2"
 FT Region 95-102
 FT /note= "grafted murine CDR3"
 PN W09201059-A.
 PD 23-JAN-1992.
 PF 05-JUL-1991; G01108.
 PR 05-JUL-1990; GB-014932.
 PR 21-DEC-1990; WD-G02017.
 PR 05-JUL-1991; WD-G01108.
 PA (CELL-) CELLTECH LTD.
 PI Adair JR, Bodmer MW, Mountain A, Owens RJ;
 DR WPI; 92-284316/34.
 DR N-PSDB; 027354.
 PT Humanised antibody molecules - comprising murine and human regions,
 PT specific for carcino-embryonic antigen, useful for diagnosis and
 PT therapy
 PS Example 4; Figure 10; 71pp; English.
 CC This sequence is CDR-grafted A5B7 human antibody having
 CC murine CDRs at amino acids 26-35 (CDR1), 50-65 (CDR2), and 95-102

CC (CDK3) and additional murine framework residues at 1, 48, 49, 72,
CC 73, 76, and 93. The LAY framework was chosen when making the coding
CC construct (Q27354) as it shows the highest homology to A5B7. The
CC antibody has specificity for carcinoembryonic antigen, produced by
CC tumours, and the Ab is thus useful in both therapy and diagnosis of
CC certain carcinomas.

Sequence 146 AA;
7 A; 7 R; 4 N; 4 D; 0 B; 2 C; 6 Q; 6 E; 0 Z; 18 G; 1 H;
3 I; 14 L; 6 K; 3 M; 9 F; 3 P; 16 S; 13 T; 5 W; 9 Y; 10 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                |||||
QAPGKGLEWLGFIGNKANGYTTEYSASVKGRFTISRDKSKSTLYLQMNGLQAEVSAIYYCTDRDGLRFYFDY
 60      70      80      90      100      110 X      120

WGQGTLLVTVSSASTKGP
130      140
```

4. US-08-249-182-4 (1-5)

R20793 CDR-grafted, humanised heavy chain gH1.

ID R20793 standard; Protein; 146 AA.
AC R20793;
DT 19-MAY-1992 (first entry)
DE CDR-grafted, humanised heavy chain gH1.
KW murine monoclonal antibody; MAb; A5B7; humanised antibody; CEA;
KW complementarity determining region.
OS Homo sapiens.
OS Mus musculus.
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal
FT Protein 20..146
FT /label= VH
FT /note= "human LAY framework with A5B7 CDRs"
FT Region 45..54
FT /label= CDR1
FT /note= "murine residues"
FT Region 69..87
FT /label= CDR2
FT /note= "murine residues"
FT Region 120..129
FT /label= CDR3
FT /note= "murine residues"
FT Misc_difference 20
FT /note= "murine residue"
FT Misc_difference 67..68
FT /note= "murine residues"
FT Misc_difference 94..95
FT /note= "murine residues"
FT Misc_difference 98
FT /note= "murine residue"
FT Misc_difference 118
FT /note= "murine residue"
PN WD9201059-A.
PD 23-JAN-1992.
PF 05-JUL-1991; G01108.
PR 05-JUL-1990; GB-014932.
PR 21-DEC-1990; WD-G02017.
PR 05-JUL-1991; WD-G01108.

PA (CELL-7) CELLTECH LTD.
 PI Adair JR, Bodmer MW, Mountain A, Owens RJ;
 DR WPI; 92-056874/07.
 DR N-PSDB; Q20987.
 PT New CDR-grafted anti carcinoembryonic antigen antibodies - useful
 PT in therapy and diagnosis of carcinoma
 PS Claim 14; Fig 10; 70pp; English.
 CC This heavy chain sequence comprises a human framework (i.e. the LAY
 CC region) which contains murine sequences (from the murine anti-CEA
 CC A5B7 MAb) in the CDRs and at other positions predicted to be
 CC important for antigen-binding and at which human and A5B7 sequences
 CC differ. (See Q20984 for A5B7 heavy chain coding sequence).
 SQ Sequence 146 AA:
 SQ 7 A; 7 R; 4 N; 4 D; 0 B; 2 C; 6 Q; 6 E; 0 Z; 18 G; 1 H;
 SQ 3 I; 14 L; 6 K; 3 M; 9 F; 3 P; 16 S; 13 T; 5 W; 9 Y; 10 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 |||||
 QAPGKGLEWLGFIGNKANGYTTSEVSASVKGRFTISRDKSKSTLYLQMNGLQAEVSAIYYCTDRGLRFYFDY
 60 70 80 90 100 110 X 120
 WGQGT LVT VSSASTKGP
 130 140

5. US-08-249-182-4 (1-5)
 R41469 MAb 25D2 humanised heavy chain.

ID R41469 standard; Protein; 140 AA.
 AC R41469;
 DT 03-MAR-1994 (first entry)
 DE MAb 25D2 humanised heavy chain.
 KW Heavy; VH; light; VL; chain; variable region; antihuman; interleukin-4;
 KW IL-4; monoclonal antibody; MAb; 25D2; single chain binding protein;
 KW complementarity determining region; CDR; humanised; Fv region; BABS;
 KW antagonist; polymerase chain reaction; PCR; primer; amplify; gamma4;
 KW pSV.SPORT.
 OS Rattus rattus.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /note= "Leader sequence"
 FT Protein 20..140
 FT /note= "25D2 H chain"
 PN W09317106-A.
 PD 02-SEP-1993.
 PF 18-FEB-1993; U01301.
 PR 19-FEB-1992; US-841659.
 PA (SCHE) SCHERING CORP.
 PI Abrams JS, Dalie B, Le HV, Miller K, Murgolo NJ;
 PI Nguyen H, Pearce M, Tindall S, Zavodny PJ;
 DR WPI; 93-288412/36.
 DR N-PSDB; Q48350.
 PT Monoclonal antibodies against human interleukin-4 corresp. DNA
 PT and CDRs - are useful for detection of interleukin-4 and treatment
 PT of related diseases
 PS Example 9; Page 89-90; 114pp; English.
 CC This sequence represents the humanised heavy (H) chain of the antihuman
 CC interleukin-4 (IL-4) monoclonal antibody (MAb) 25D2. The 25D2 H
 CC chain coding region was cloned in three fragments using the primers
 CC given in Q48351-60. The amplified fragments were designed to

CC contain silent restriction sites, however several codons had to
 CC be changed to incorporate further restriction sites. The primers
 CC given in Q48361-66 were used to amplify the entire H chain variable
 CC region (VH) of an unrelated humanised antibody. The amplified fragments
 CC were then cloned into pSV.Sport which already contained the 25D2 H
 CC chain fragments. The primers given in Q48367-72 were used in
 CC further manipulations to amplify a human gamma4 constant region cDNA
 CC which was used to replace the genomic DNA. The humanised MAb is an
 CC IL-4 antagonist. It may be used in a pharmaceutical composition for
 CC detecting, measuring and immunopurifying human IL-4 and blocking IL-4
 CC activity in IL-4-related diseases.

SO Sequence 140 AA;

SO 9 A; 7 R; 4 N; 5 D; 0 B; 3 C; 6 Q; 4 E; 0 Z; 16 G; 2 H;

SO 6 I; 12 L; 2 K; 3 M; 7 F; 4 P; 18 S; 8 T; 5 W; 9 Y; 10 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                |||||
VRQAPGKGLEWVASISISGDNITYYPDSVRGRFTISRND SKNTLYLQMNGLQAEVSAIYYCARDPYFSGHYF
    60      70      80      90      100      X 110      120

DFWGQGTLTVTSS
    130      140

```

6. US-08-249-182-4 (1-5)

R40179 Humanised antibody CMX5-3 heavy chain variable reg

ID R40179 standard; Protein; 135 AA.

AC R40179;

DT 14-FEB-1994 (first entry)

DE Humanised antibody CMX5-3 heavy chain variable region.

KW Primer; polymerase chain reaction; amplify; PCR; human; kappa; L;

KW constant region; heavy; H; chain; pUC19; humanised; antibody;

KW light; REI; VL3 fragment; CMX5-1; CMX5-3.

OS Synthetic.

FH Key Location/Qualifiers

FT Peptide 1..19

FT /note= "Secretory leader peptide"

FT Protein 20..135

FT /note= "CMX5-3 heavy chain variable region"

PN W09316184-A.

PD 19-AUG-1993.

PF 04-FEB-1993; U00759.

PR 06-FEB-1992; US-832842.

PA (SCHE) SCHERING CORP.

PI Abrams JS, Chou C, Jenh C, Murgolo NJ, Petro ME;

PI Silver JE, Tindall S, Windsor WT, Zavodny PJ;

DR WPI; 93-272888/34.

PT Humanised monoclonal antibody - comprises variable animal region

PT and constant human region, binds to human interleukin-5

PS Example; Page 91-92; 118pp; English.

CC The sequences given in R40179-80 represent the variable regions of

CC the heavy and light chains of the humanised antibody CMX5-3

CC respectively. These sequences were based on the humanised antibody

CC CMX5-1. These sequences were generated using the primer sequences

CC given in Q48068-71. These primers were based on sequences derived

CC from antibody JES1-39D10 and human LAY VH framework sequences. The

CC amplification products were used to replace the VH1 and VH3 fragments

CC of CMX5-1 H chain cDNA in pSV.Sport (see also R40175).

SO Sequence 135 AA;

SO 9 A; 5 R; 7 N; 4 D; 0 B; 3 C; 6 Q; 5 E; 0 Z; 15 G; 1 H;

Initial Score = 5 Optimized Score = 5 Significance = 4.48
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                @AEVS
                                |||||
WIRQAPGKGLEWVALIWSNGDIDYNSAIKSRFTISRNDKNTLYLQMNGL@AEVSAIYFCAREYYGYFDYWG
    60      70      80      90      100  X  110      120

GGTLVTVSS
130

```

7. US-08-249-182-4 (1-5)

R41113 HCV peptide XIV or HCV8 (aa 1730-1749).

ID R41113 standard; peptide; 22 AA.
 AC R41113;
 DT 22-MAR-1994 (first entry)
 DE HCV peptide XIV or HCV8 (aa 1730-1749).
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified_site 1
 FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where
 FT B= biotin;
 FT X= biotinylation cpd. incorporated
 FT during synthesis;
 FT Y= bond or linking gp(s). which
 FT minimises steric hindrance,
 FT where Y is not a bond it is pref. 1-10
 FT residues of (same or different) glycine,
 FT beta-alanine, 4-aminobutyric acid,
 FT 5-aminovaleric acid or 6-aminohexanoic acid;
 FT parenthesis around B and X indicate opt. presence
 FT at the specified positions but B or X must be
 FT present in at least one of the positions shown,
 FT B interacts with the peptide to give a cpd.
 FT with greater diagnostic sensitivity;
 FT A (optional)= one or more amino acids, NH2 or
 FT gp. which modifies the N-terminus;
 FT Z= one or more amino acids, OH, NH2, or a
 FT linkage involving either of these 2 gps."
 FT Modified_site 22
 FT /note= "the C-terminal comprises Y-(X)-Z"
 PN W09318054-A.
 PD 16-SEP-1993.
 PF 08-MAR-1993; E00517.
 PR 06-MAR-1992; EP-400598.
 PA (INNO-) INNOGENETICS NV SA.
 PI De LEYS R;
 DR WPI; 93-303397/38.
 PT New biotinylated peptide(s) corresp. to immuno-dominant
 PT epitope(s) - with increased antigenicity, useful in antibodies
 PT detection and vaccines against hepatitis C, HIV and HTLV
 PS Claim 4; Page 90-98; 133pp; English.
 CC Peptide compsns. comprise at least one and pref. a combination of
 CC two, three, four or more biotinylated peptides chosen from the
 CC sequences given in R41058-R41166. The peptides represent
 CC immunologically important regions of viral proteins and are
 CC prepd. by solid phase peptide synthesis. The compsns. are

CC Useful for the detection of antibodies to HCV, and/or HIV,
CC and/or HTLV-I or II.
SQ Sequence 22 AA;
SQ 5 A; 1 R; 0 N; 0 D; 0 B; 0 C; 3 Q; 1 E; 0 Z; 1 G; 0 H;
SQ 1 I; 3 L; 1 K; 0 M; 0 F; 1 P; 1 S; 1 T; 0 W; 0 Y; 1 V;
SQ 2 Others;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||
XQKALGLLQTASRQAEVIAPAX
10 X 20

8. US-08-249-182-4 (1-5)

R41186 HCV NS4 protein HCV7/8.

ID R41186 standard; peptide; 26 AA.
AC R41186;
DT 22-MAR-1994 (first entry)
DE HCV NS4 protein HCV7/8.
KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
OS Synthetic.
FH Key Location/Qualifiers
FT Region 10..16
FT /label= epitope_4
PN WD9318054-A.
PD 16-SEP-1993.
PF 08-MAR-1993; E00517.
PR 06-MAR-1992; EP-400598.
PA (INNO-) INNOGENETICS NV SA.
PI De LEYS R;
DR WPI; 93-303397/38.
PT New biotinylated peptide(s) corresp. to immuno-dominant
PT epitope(s) - with increased antigenicity, useful in antibodies
PT detection and vaccines against hepatitis C, HIV and HTLV
PS Disclosure; Page 79; 133pp; English.
CC Peptide compsns. comprise at least one and pref. a combination of
CC two, three, four or more biotinylated peptides chosen from the
CC sequences given in R41058-R41166.
CC The peptides may be hybrids consisting of combinations of the core
CC epitopes of the HCV core (R41171-R41180), HCV NS4 (R41181-R41186) or
CC the HCV NS5 (R41187-R41193) region separated by Gly and/or Ser residues.
CC Pref. hybrid peptides are given in R41161-R41163.
CC The peptides represent immunologically important regions of viral
CC proteins and are prepd. by solid phase peptide synthesis. The compsns.
CC are useful for the detection of antibodies to HCV, and/or HIV,
CC and/or HTLV-I or II.
SQ Sequence 26 AA;
SQ 6 A; 1 R; 0 N; 0 D; 0 B; 0 C; 4 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 4 L; 2 K; 0 M; 1 F; 1 P; 1 S; 1 T; 0 W; 0 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||
LAEQFKQKALGLLQTASRQAEVIAPA

9. US-08-249-182-4 (1-5)

P90810 Synthetic HTLV-1 peptide antigen.

ID P90810 standard; peptide; 26 BP.
 AC P90810;
 DT 07-FEB-1990 (first entry)
 DE Synthetic HTLV-1 peptide antigen.
 KW HTLV-1; antigen; vaccine; env gene.
 FH Key Location/Qualifiers
 FT Region 1
 FT Region 25
 FT Region 26
 PN W08908664-A.
 PD 21-SEP-1989.
 PF 10-MAR-1989; SE0126.
 PR 10-MAR-1988; US-166205.
 PA (VIRO-) Virovahl S.A.
 PI Svennerholm B, Rymo L, Jeansson S, Horal P;
 DR WPI; 89-292495/40.
 PT Synthetic HTLV-1 peptide antigens - used for detection of HTLV-1
 PT infection or in vaccines to elicit prodn. of antibodies.
 PS Claim 3; page 25; 34pp; English.
 CC The synthetic peptide sequence corresponds to an immunodominant region
 CC of the envelope glycoprotein encoded by bps 6018-6086 of the env gene of
 CC HTLV-1. AA1 = H or an amino acid added to facilitate coupling to a
 CC carrier protein; AA25 = absent or Cys; and AA26 = OH or NH2. The
 CC Ag peptide is used for detection of HTLV-1 infection or in vaccines
 CC See also P90809, P90811, and P90812.
 SQ Sequence 26 AA;
 SQ 1 A; 0 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 0 G; 2 H;
 SQ 3 I; 2 L; 0 K; 0 M; 1 F; 2 P; 3 S; 1 T; 1 W; 0 Y; 1 V;
 SQ 3 Others;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 || ||
 XWTHCFDPQIGAIQVSSPCHNSLILXX
 10 X 20

10. US-08-249-182-4 (1-5)

R40053 Hib OMP P1 peptide HIBP1-1 (1-29).

ID R40053 standard; peptide; 30 AA.
 AC R40053;
 DT 04-FEB-1994 (first entry)
 DE Hib OMP P1 peptide HIBP1-1 (1-29).
 KW Haemophilus influenzae; type b; Hib; outer membrane protein; P1; P2;
 KW P6; vaccine; antibody; detection; lipoglycopeptide conjugate;
 KW immunogen.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 30
 FT /note= "May be absent"
 PN W09315205-A.
 PD 05-AUG-1993.
 PF 03-FEB-1993; CA0041.
 PR 03-FEB-1992; GB-002219.
 PA (CONN-) CONNAUGHT LAB LTD.

PI Chong P, Kandil A, Klein HH, Sia C;
 DR WPI; 93-258681/32.
 PT Synthetic Haemophilus influenzae conjugate vaccine - comprising
 PT T-helper cell determinants and B-cell epitope(s) linked to
 PT synthetic oligo:saccharide(s)
 PS Table 1; Page 47; 99pp; English.
 CC The sequences given in R40053-101 are peptide fragments derived from
 CC the Haemophilus influenzae type b (Hib) outer membrane proteins P1,
 CC P2 and P6. These peptides may be used in a vaccine against Hib
 CC infection and antibodies against these peptides may be used in test
 CC kits to detect H. influenzae in a sample. The vaccine may further
 CC comprise a immunogenic or immunostimulatory molecule or the peptides
 CC may be modified with lipids, or linked to synthetic PRP as synthetic
 CC lipoglycopeptide conjugates to produce alternative vaccines.
 SQ Sequence 30 AA;
 SQ 9 A; 1 R; 1 N; 1 D; 0 B; 1 C; 1 Q; 2 E; 0 Z; 3 G; 0 H;
 SQ 1 I; 2 L; 0 K; 0 M; 1 F; 0 P; 3 S; 1 T; 0 W; 1 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 AAFQLAEVSTSGGLGRAYAGEAAIADNASVC
 X 10 20 30

11. US-08-249-182-4 (1-5)
 R13186 Peptide VII immunoreactive with anti-HTLV antibody

ID R13186 standard; Protein; 38 AA.
 AC R13186;
 DT 09-OCT-1991 (first entry)
 DE Peptide VII immunoreactive with anti-HTLV antibodies.
 KW human T-cell leukaemia virus; AIDS; ATL; detection;
 KW envelope protein gp61; acquired immunodeficiency syndrome.
 OS Synthetic.
 PN EP-439077-A.
 PD 31-JUL-1991.
 PF 18-JAN-1991; 100616.
 PR 24-JAN-1990; US-469291.
 PA (UNBI-) UNITED BIOMEDICAL.
 PI Wang CY;
 DR WPI; 91-224505/31.
 PT Peptide compsns. corresp. to envelope fragments of HTLV-1,2 - for
 PT detecting antibodies to these viruses and diagnosing HIV and
 PT adult T-cell leukaemia infections
 PS Claim 1; Page 17; 27pp; English.
 CC This peptide is one of 16 peptides useful for detecting antibodies to
 CC HTLV or HIV viruses. The peptides correspond to partial sequences of
 CC the HTLV virus designated gp21 and gp64, both part of gp61, which
 CC defines the envelope protein of the HTLV-I or HTLV-II virus. The
 CC peptides can be amidated at the C-terminal. This particular peptide
 CC is used in a composition with at least two of the other peptides of
 CC the invention. See R13184-R13193 and R13861-6.
 SQ Sequence 38 AA;
 SQ 2 A; 3 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 3 I; 4 L; 0 K; 0 M; 2 F; 6 P; 7 S; 1 T; 0 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X

QAEVS
 || ||
 CFDPQIQAIIVSSPCHNSLILPPFSLSPVPTLGSRSRRA
 X 10 20 30

12. US-08-249-182-4 (1-5)

R06408 HTLV-1 corresponding peptide (VI).

ID R06408 standard; protein; 38 AA.
 AC R06408;
 DT 21-DEC-1990 (first entry)
 DE HTLV-1 corresponding peptide (VI).
 KW HTLV-1; HIV; antibodies; vaccines; polymers;
 OS Synthetic.
 PN W09008162-A.
 PD 26-JUL-1990.
 PF 16-JAN-1990; U00260.
 PR 13-JAN-1989; US-297635.
 PA (UNBI-) UNITED BIOMED INC.
 PI Yang CY;
 DR WPI; 90-254015/33.
 PT Synthetic peptide(s) corresponding to HTLV-1 and op. HIV - used
 PT for detection of antibodies, in vaccines and for development of
 PT antibodies
 PS Claim 1 (VI); Page 38; 52pp; English.
 CC Peptides having specific immunoreactivity to antibodies to HTLV-1
 CC comprise this sequence on its own, or an analogue of it in which
 CC amino acids may be added, deleted or substd, or segments, mixts.,
 CC conjugates or polymers of the peptides representes in R06403-08.
 CC The peptides are safe, sensitive and specific in the detection of
 CC antibodies. This peptide corresponds to a partial segment of the
 CC amino acid sequence of the HTLV-1 virus gp.21 or gp.46 and are
 CC prepared by solid phase synthesis.
 SQ Sequence 38 AA;
 SQ 2 A; 3 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 3 I; 4 L; 0 K; 0 M; 2 F; 6 P; 7 S; 1 T; 0 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 || ||
 CFDPQIQAIIVSSPCHNSLILPPFSLSPVPTLGSRSRRA
 X 10 20 30

13. US-08-249-182-4 (1-5)

R20772 Peptide V based on immunoreactive region of Hepati

ID R20772 standard; Protein; 40 AA.
 AC R20772;
 DT 05-MAY-1992 (first entry)
 DE Peptide V based on immunoreactive region of Hepatitis C virus.
 KW Non-A, non-B hepatitis virus; non-structural protein; vaccine.
 OS Synthetic.
 PN EP-468527-A.
 PD 29-JAN-1992.
 PF 26-JUL-1991; 112620.
 PR 26-JUL-1990; US-558799.
 PR 07-FEB-1991; US-651735.
 PR 11-MAR-1991; US-667275.
 PR 24-JUN-1991; US-719819.
 PA (UNBI-) UTD BIOMEDICAL INC.

FI Chang TW; Husein B;
 DR WPI: 92-034279/05.
 PT New synthetic peptide specific for HCV antibodies - for detection
 PT of HCV or NANBH e.g. by enzyme-linked immunosorbent assay and is
 PT immunogen for preparation of vaccines
 PS Disclosure: Page 13; 98pp; English.
 CC Peptide V is from the non-structural protein region of HCV. It was
 CC found to be reactive and useful for the detection of antibodies to
 CC HCV and diagnosis of NANBH. See R20751-R20782.
 SQ Sequence 40 AA;
 SQ 6 A; 1 R; 2 N; 0 D; 0 B; 0 C; 5 Q; 2 E; 0 Z; 1 G; 1 H;
 SQ 1 I; 4 L; 4 K; 1 M; 2 F; 1 P; 1 S; 3 T; 3 W; 0 Y; 2 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 KQKALGLLQTASRQAEVIAPAVQTNWQKLETFWAKHMMWF
 10 X 20 30 40

14. US-08-249-182-4 (1-5)
 R13185 Peptide (V) immunoreactive with anti-HTLV antibody

ID R13185 standard; Protein; 40 AA.
 AC R13185;
 DT 09-OCT-1991 (first entry)
 DE Peptide (V) immunoreactive with anti-HTLV antibodies.
 KW human T-cell leukaemia virus; AIDS; ATL; detection;
 KW envelope protein gp61; acquired immunodeficiency syndrome.
 OS Synthetic.
 PN EP-439077-A.
 PD 31-JUL-1991.
 PF 18-JAN-1991; 100616.
 PR 24-JAN-1990; US-469291.
 PA (UNBI-) UNITED BIOMEDICAL.
 PI Wang CY;
 DR WPI: 91-224505/31.
 PT Peptide compsns. corresp. to envelope fragments of HTLV-1,2 - for
 PT detecting antibodies to these viruses and diagnosing HIV and
 PT adult T-cell leukaemia infections
 PS Claim 1; Page 17; 27pp; English.
 CC This peptide is one of 16 peptides useful for detecting antibodies to
 CC HTLV or HIV viruses. The peptides correspond to partial sequences of
 CC the HTLV virus designated gp21 and gp64, both part of gp61, which
 CC defines the envelope protein of the HTLV-I or HTLV-II virus. The
 CC peptides can be amidated at the C-terminal. This particular peptide
 CC is used in a composition with at least two of the other peptides of
 CC the invention. See also R13184, R13186-R13192 and R13861-6.
 SQ Sequence 40 AA;
 SQ 3 A; 0 R; 1 N; 1 D; 0 B; 2 C; 2 Q; 0 E; 0 Z; 0 G; 3 H;
 SQ 2 I; 6 L; 0 K; 0 M; 2 F; 7 P; 5 S; 3 T; 1 W; 1 Y; 1 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 SSTPLLYPSLALPAPHLTLFPNWTCHCFDPQIGAIIVSSPCH
 10 20 30 X X 40

15. US-08-249-182-4 (1-5)
R13591 HTLV-I env precursor epitope peptide.

ID R13591 standard; Protein; 40 AA.
AC R13591;
DT 03-OCT-1991 (first entry)
DE HTLV-I env precursor epitope peptide.
KW HTLV-I; epitope; diagnosis; env protein; gp46; antibody; vaccine.
OS Synthetic.
PN US5003043-A.
PD 26-MAR-1991.
PF 25-MAY-1988; 198416.
PR 25-MAY-1988; US-198416.
PA (TRIT-) TRITON BIOSCIENCES.
PI Akita RW, Florine DL, Ralston JS;
DR WPI; 91-221557/30.
PT Synthetic peptide(s) and antibodies corresp. to an epitope of
PT HTLV-I - used in diagnosis, therapy prepn. of vaccines and
PT prognostic indicators of HTLV-I infection
PS Disclosure; Page 3; 10pp; English.
CC The peptide has specific binding affinity for O.5alpha monoclonal
CC antibody. It represents an epitopic site on the major HTLV-I
CC envelope precursor, i.e it corresponds to residues 281-320. The
CC peptide and its antibody can be used in diagnosing the presence of
CC HTLV-I associated diseases, as vaccines against HTLV-I infection or
CC as prognostic indicators after HTLV-I infection.
CC See also R13077, R13590 and R13591.
SQ Sequence 40 AA;
SQ 3 A; 3 R; 0 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 1 G; 1 H;
SQ 3 I; 5 L; 0 K; 0 M; 1 F; 6 P; 7 S; 1 T; 1 W; 0 Y; 5 V;

Initial Score = 4 Optimized Score = 4 Significance = 3.58
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

      X  X
      QAEVS
      || ||
      IQAIVSSPCHCSLILPPFSLSPVPTLGSRSRRRAVPVAVWL
      X  X 10      20      30      40
> D <
0| |0 IntelliGenetics
> D <
```

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

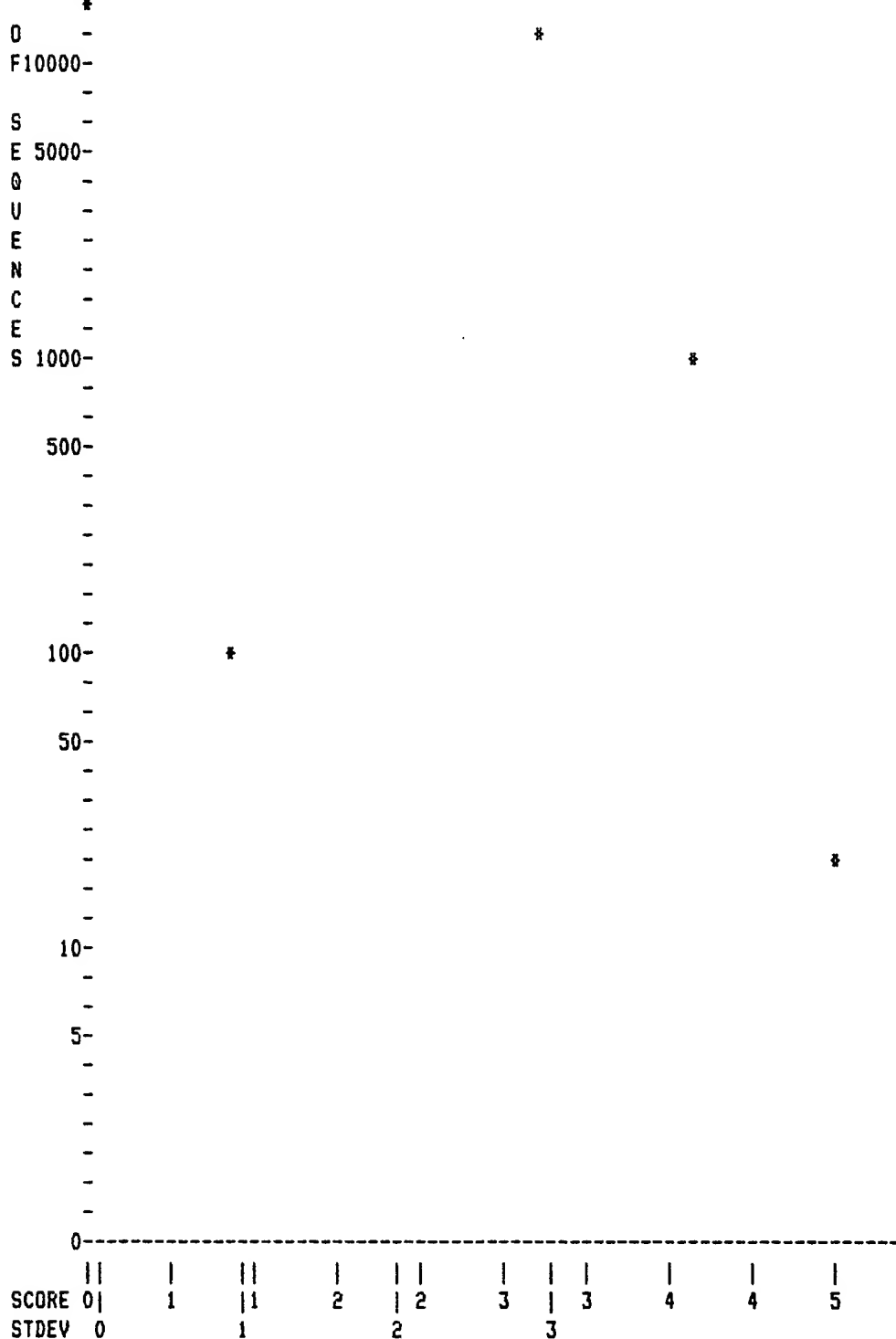
Results file u249_4p.res made by on Thu 22 Sep 94 10:25:41-PDT.

Query sequence being compared:US-08-249-182-4 (1-5)
Number of sequences searched: 70848
Number of scores above cutoff: 4444

Results of the initial comparison of US-08-249-182-4 (1-5) with:
Data bank : PIR 41, all entries

100000-
-
N -
U50000-
M -
B -
E -
R -

*



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.95

Times:	CPU	Total Elapsed

Number of residues: 20816057
 Number of sequences searched: 70848
 Number of scores above cutoff: 4444

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

19 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Opt.		Sig.	Frame
			Score	Score		
1. S20590	sialidase (EC 3.2.1.18) - Act	913	5	5	4.19	0
2. A49227	sialidase - Actinomyces visco	901	5	5	4.19	0
3. S39558	HSP90 homolog - Madagascar pe	817	5	5	4.19	0
4. S33533	heat shock protein 90 homolog	809	5	5	4.19	0
5. S31862	GRP94 protein homolog - barle	809	5	5	4.19	0
6. A33827	regulatory protein ral2 - fis	611	5	5	4.19	0
7. WMCVFM	inclusion body matrix protein	512	5	5	4.19	0
8. A30828	steroid 17alpha-monooxygenase	507	5	5	4.19	0
9. S16719	steroid 17alpha-monooxygenase	507	5	5	4.19	0
10. S38397	vasoactive intestinal peptide	460	5	5	4.19	0
11. JH0594	vasoactive intestinal peptide	459	5	5	4.19	0
12. JN0604	vasoactive intestinal peptide	457	5	5	4.19	0
13. S16562	nolF protein - Rhizobium meli	367	5	5	4.19	0
14. A48470	elongation factor 1 alpha, EF	346	5	5	4.19	0
15. A27659	cytochrome P450 17 - rat (fra	237	5	5	4.19	0
16. A33980	steroid 17alpha-monooxygenase	235	5	5	4.19	0
17. S18659	hypothetical protein - Mycopl	148	5	5	4.19	0
18. S06727	hypothetical protein 1 (minic	122	5	5	4.19	0
19. A42329	autotaxin - human (fragments)	114	5	5	4.19	0

The list of other best scores is:

		Init. Opt.				
Sequence Name	Description	Length	Score	Score	Sig.	Frame

**** 3 standard deviations above mean ****						
20. A35776	RecQ protein - Escherichia co	5	4	4	3.14	0
21. PC1149	equinatoxin 1A - sea anemone	13	4	4	3.14	0
22. A40634	orf19 3' of eryK - Saccharopo	15	4	4	3.14	0
23. S35970	ribosomal protein L10 - Citro	20	4	4	3.14	0
24. S35978	ribosomal protein L10 - Prote	21	4	4	3.14	0
25. B60701	31K antigen - Campylobacter j	22	4	4	3.14	0
26. S35976	ribosomal protein L10 - Klebs	23	4	4	3.14	0
27. S35975	ribosomal protein L10 - Enter	23	4	4	3.14	0
28. S04171	aadA protein - Klebsiella pne	23	4	4	3.14	0
29. A60701	31K antigen PEB4 - Campylobac	38	4	4	3.14	0
30. S10765	glutamate synthase - Azospiri	40	4	4	3.14	0
31. A10265	alpha-lactalbumin I - eastern	42	4	4	3.14	0
32. B39880	streptomycin/spectinomycin re	42	4	4	3.14	0
33. S22426	thymosin beta-12 - rainbow tr	44	4	4	3.14	0
34. S25708	hypothetical protein 4a - hum	44	4	4	3.14	0
35. S11162	photosystem II protein psbK p	46	4	4	3.14	0
36. A05024	hypothetical protein 55 - liv	55	4	4	3.14	0
37. S01585	photosystem II protein psbK p	55	4	4	3.14	0
38. S30963	gene 18 protein - Mycobacteri	57	4	4	3.14	0
39. S31447	photosystem II protein psbK -	61	4	4	3.14	0
40. S28768	photosystem II protein psbK -	61	4	4	3.14	0

do you need
to see?

1. US-08-249-182-4 (1-5)

S20590 sialidase (EC 3.2.1.18) - *Actinomyces viscosus*

ENTRY S20590 #type complete
 TITLE sialidase (EC 3.2.1.18) - *Actinomyces viscosus*
 ORGANISM #formal_name *Actinomyces viscosus*
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993
 ACCESSIONS S20590
 REFERENCE S20590
 #authors Henningsen, M.; Roggentin, P.; Schauer, R.
 #journal Biol. Chem. Hoppe-Seyler (1991) 372:1065-1072
 #title Cloning, sequencing and expression of the sialidase gene from *Actinomyces viscosus* DSM 43798.
 #cross-references MUID:92162190
 #accession S20590
 ##status preliminary
 ##residues 1-913 ##label HEN
 ##cross-references EMBL:X62276
 SUMMARY #length 913 #molecular-weight 96216 #checksum 4303
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 |||||

TETGKKVGYSDPSYVVVDHQTGTIFNFHVKSVDQGWGSGRGCTDPENRGIIQAEVSTSTDNGWTWTHRTITAD
 370 380 390 400 410 420 X 430 440

ITKDKPWATARFAASGQGIQIHGPHAGRLVQQY
 450 460 470

2. US-08-249-182-4 (1-5)

A49227 sialidase - *Actinomyces viscosus*

ENTRY A49227 #type complete
 TITLE sialidase - *Actinomyces viscosus*
 ORGANISM #formal_name *Actinomyces viscosus*
 DATE 19-Dec-1993; #sequence_revision 19-Dec-1993; #text_change 19-Dec-1993
 ACCESSIONS A49227
 REFERENCE A49227
 #authors Yeung, M.K.
 #journal Infect. Immun. (1993) 61:109-116
 #title Complete nucleotide sequence of the *Actinomyces viscosus* T14V sialidase gene: presence of a conserved repeating sequence among strains of *Actinomyces* spp.
 #cross-references MUID:93114861
 #contents T14V
 #accession A49227
 ##status preliminary
 ##molecule_type nucleic acid
 ##residues 1-901 ##label YEU
 ##cross-references NCBIN:121598; NCBIP:121599
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 901 #molecular-weight 92860 #checksum 7681
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0

gaps - conservative substitutions -

X X
QAEVS
||||

TETGKKVGYSDPSYVVDHQTGTIFNFHVKSVDQGWGGSRGGTDPENRGIIQAEVSTSTDNGWTWTHRTITAD
370 380 390 400 410 420 X 430 440

ITKDKPWITARFAASGGGIQIQHGPAGRLVQQY
450 460 470

3. US-08-249-182-4 (1-5)

S39558 HSP90 homolog - Madagascar periwinkle

ENTRY S39558 #type complete
TITLE HSP90 homolog - Madagascar periwinkle
ORGANISM #formal_name Catharanthus roseus #common_name Madagascar periwinkle
DATE 19-May-1994; #sequence_revision 19-May-1994; #text_change 19-May-1994
ACCESSIONS S39558
REFERENCE S39558
#authors Schroeder, G.; Beck, M.; Eichel, J.; Vetter, H.P.; Schroeder, J.
#journal Plant Mol. Biol. (1993) 23:583-594
#title HSP90 homologue from Madagascar periwinkle (Catharanthus roseus): cDNA sequence, regulation of protein expression and location in the endoplasmic reticulum.
#accession S39558
##status preliminary
##residues 1-817 ##label SCH
SUMMARY #length 817 #molecular-weight 93491 #checksum 9348
SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||

DSDAPVDPKPKVEDKIGAVPNGLSTDSDVAKREAESMSMRNLRSDAEKFEFQAEVSRMLMDIINSLYSNKIDIF
40 50 60 70 80 X X 90 100

LRELISNASDALDKIRFLALTDKEILGEGDTAK
110 120 130

4. US-08-249-182-4 (1-5)

S33533 heat shock protein 90 homolog precursor - barley

ENTRY S33533 #type complete
TITLE heat shock protein 90 homolog precursor - barley
ORGANISM #formal_name Hordeum vulgare #common_name barley
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-May-1994
ACCESSIONS S33533
REFERENCE S33533
#authors Walther-Larsen, H.; Brandt, J.; Collinge, D.B.; Thordal-Christensen, H.
#journal Plant Mol. Biol. (1993) 21:1097-1108
#title A pathogen-induced gene of barley encodes a HSP90 homologue showing striking similarity to vertebrate forms resident in the endoplasmic reticulum.
#accession S33533


```

##molecule_type mRNA
##residues 1-809 ##label WAL
##cross-references EMBL:X67960
CLASSIFICATION #superfamily heat shock protein 90
KEYWORDS glycoprotein
FEATURE
  1-20 #domain signal sequence #status predicted #label SIG\
  21-809 #protein heat shock protein 90 homolog #status predicted
        #label MAT\
  111,410,450,617 #binding_site carbohydrate (Asn) (covalent) #status
                  predicted
SUMMARY #length 809 #molecular-weight 92916 #checksum 2897
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.19
Residue Identity   =  100%  Matches          =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

                                X  X
                                QAEVS
                                |||||
SSDEVGDFPKVEEKLGA VPHGLSTDSEVVQRESESI SRKTLRNSAEKFEFQAEVSRLMDIIINSLYSNKDIF
  40      50      60      70      80 X  X 90      100

LRELISNASDALDKIRFLALTDKEVMGEGDTAK
  110      120      130

```

5. US-08-249-182-4 (1-5)

S31862 GRP94 protein homolog - barley

```

ENTRY      S31862      #type complete
TITLE      GRP94 protein homolog - barley
ORGANISM   #formal_name Hordeum vulgare #common_name barley
DATE       20-May-1994; #sequence_revision 20-May-1994; #text_change
           20-May-1994
ACCESSIONS S31862
REFERENCE  S31862
  #authors  Brandt, J.
  #submission submitted to the EMBL Data Library, August 1992
  #accession S31862
    ##status preliminary
    ##residues 1-809 ##label BRA
    ##cross-references EMBL:X67960
SUMMARY    #length 809 #molecular-weight 92916 #checksum 2897
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.19
Residue Identity   =  100%  Matches          =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

                                X  X
                                QAEVS
                                |||||
SSDEVGDFPKVEEKLGA VPHGLSTDSEVVQRESESI SRKTLRNSAEKFEFQAEVSRLMDIIINSLYSNKDIF
  40      50      60      70      80 X  X 90      100

LRELISNASDALDKIRFLALTDKEVMGEGDTAK
  110      120      130

```

6. US-08-249-182-4 (1-5)

A33827 regulatory protein ral2 - fission yeast (Schizosac

```

ENTRY      A33827      #type complete
TITLE      regulatory protein ral2 - fission yeast (Schizosaccharomyces

```

ORGANISM #formal_name Schizosaccharomyces pombe
 DATE 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change
 30-Sep-1993
 ACCESSIONS A33827
 REFERENCE A33827
 #authors Fukui, Y.; Miyake, S.; Satoh, M.; Yanamoto, M.
 #journal Mol. Cell. Biol. (1989) 9:5617-5622
 #title Characterization of the Schizosaccharomyces pombe ral2 gene
 implicated in activation of the ras1 gene product.
 #cross-references MUID:90066514
 #accession A33827
 ##status preliminary
 ##molecule_type DNA
 ##residues 1-611 ##label FUK
 ##cross-references GB:M30827
 SUMMARY #length 611 #molecular-weight 69847 #checksum 9734
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 @AEVS
 ||||
 GKLLNGISDMEILTIERMHIPCLSRMLYKRWPAFQKIMDRAVEKNQAEF@AEVSQLGPLTDLPFSSIHST
 430 440 450 460 470 480 490
 GSRALYMPYSFETCSAFLHYIYCGTLNGSYCTA
 500 510 520 530

7. US-08-249-182-4 (1-5)

WMCVFM inclusion body matrix protein - figwort mosaic vir

ENTRY WMCVFM #type complete
 TITLE inclusion body matrix protein - figwort mosaic virus
 ORGANISM #formal_name figwort mosaic virus
 DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
 08-Apr-1994
 ACCESSIONS S01284
 REFERENCE S01279
 #authors Richins, R.D.; Scholthof, H.B.; Shepherd, R.J.
 #journal Nucleic Acids Res. (1987) 15:8451-8466
 #title Sequence of figwort mosaic virus DNA (caulimovirus group).
 #cross-references MUID:88040466
 #accession S01284
 ##molecule_type DNA
 ##residues 1-512 ##label RIC
 ##cross-references EMBL:X06166
 ##note the translation of the nucleotide sequence is not given
 in this paper
 CLASSIFICATION #superfamily caulimovirus inclusion body matrix protein
 KEYWORDS matrix protein
 SUMMARY #length 512 #molecular-weight 58207 #checksum 1605
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 @AEVS
 ||||
 FKPDYLRASNGQSWFAVYKGPKNKEFFTEWEIVADICKKRQKSKRFRSKE@AEVSISLYNKDIQDPVNFLRP

150 160 170 180 190 X X 200 210
VKLVKEERAAQPLKFKAIAAEQTIQFDEFRQIW
220 230 240

8. US-08-249-182-4 (1-5)

A30828 steroid 17alpha-monooxygenase (EC 1.14.99.9) cytoc

ENTRY A30828 #type complete
TITLE steroid 17alpha-monooxygenase (EC 1.14.99.9) cytochrome P450
17 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 19-May-1989 #sequence_revision 19-May-1989 #text_change
28-Apr-1993
ACCESSIONS A30828; A31359
REFERENCE A94511
#authors Dufau, M.L.
#submission submitted to GenBank, December 1988
#accession A30828
##molecule_type mRNA
##residues 1-507 ##label DUF
REFERENCE A90154
#authors Naniki, M.; Kitamura, M.; Buczko, E.; Dufau, M.L.
#journal Biochem. Biophys. Res. Commun. (1988) 157:705-712
#title Rat testis P-450-17-alpha cDNA: the deduced amino acid
sequence, expression and secondary structural
configuration.
#cross-references MUID:89076306
#accession A31359
##molecule_type mRNA
##residues 1-507 ##label NAM
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS endoplasmic reticulum; heme; membrane protein; monooxygenase;
oxidoreductase
FEATURE
2-21 #domain transmembrane #label TM1\
169-186 #domain transmembrane #label TM2\
441 #binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY #length 507 #molecular-weight 57250 #checksum 9025
SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X X
QAEVS
||||
FVFTALLLRFDLDVSDDKQLPRLEGDPKVVFLIDPFVKITVRQAWMDAQAEVST
460 470 480 490 500 X X

9. US-08-249-182-4 (1-5)

S16719 steroid 17alpha-monooxygenase (EC 1.14.99.9) cytoc

ENTRY S16719 #type complete
TITLE steroid 17alpha-monooxygenase (EC 1.14.99.9) cytochrome P450
17 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change
21-Nov-1993
ACCESSIONS S16719
REFERENCE S16719
#authors Fevold, H.R.; Lorence, M.C.; McCarthy, J.L.; Trant, J.M.;

#journal Mol. Endocrinol. (1989) 3:968-975
 #title Rat P450(17-alpha) from testis: characterization of a
 full-length cDNA encoding a unique steroid hydroxylase
 capable of catalyzing both Delta(4)- and Delta
 (5)-steroid-17,20-lyase reactions.

#cross-references MUID:89295447

#accession S16719

##status preliminary

##residues 1-507 ##label FEV

##cross-references EMBL:M31681

SUMMARY #length 507 #molecular-weight 57250 #checksum 9025

SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                |||||
FVFTALLLQRFDLQVSDDKQLPRLEGDPKVVFLIDPFVKVITVRQAWMDAQAEVST
  460      470      480      490      500 X  X

```

10. US-08-249-182-4 (1-5)

S38397 vasoactive intestinal peptide receptor - human

ENTRY S38397 #type complete
 TITLE vasoactive intestinal peptide receptor - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 22-Jan-1994; #sequence_revision 22-Jan-1994; #text_change
 22-Jan-1994

ACCESSIONS S38397

REFERENCE S38397

#authors Couvineau, A.

#submission submitted to the EMBL Data Library, September 1993

#accession S38397

##status preliminary

##residues 1-460 ##label COU

##cross-references EMBL:X75299

SUMMARY #length 460 #molecular-weight 51929 #checksum 1116

SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X  X
                                QAEVS
                                |||||
RKWRRWHLQGVLGWNPKYRHPSGGSNGATCSTQVSMLTRVSPGARRSSSFQAEVSLV
  410      420      430      440      450 X  460

```

11. US-08-249-182-4 (1-5)

JH0594 vasoactive intestinal peptide receptor precursor -

ENTRY JH0594 #type complete
 TITLE vasoactive intestinal peptide receptor precursor - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
 18-Jun-1993

ACCESSIONS JH0594

REFERENCE JH0594

#authors Ishihara, T.; Shigemoto, R.; Mori, K.; Takahashi, K.; Nagata,

```

#journal      Neuron (1992) 8:811-819
#title        Functional expression and tissue distribution of a novel
               receptor for vasoactive intestinal polypeptide.
#cross-references MUID:92232309
#contents     Lung
#accession    JH0594
               ##molecule_type mRNA
               ##residues      1-459 ##label ISH
               ##cross-references GB:M86835
KEYWORDS      glycoprotein; membrane protein
FEATURE
  1-30         #domain signal sequence #status predicted #label SIG\
  31-459       #protein vasoactive intestinal polypeptide receptor
               #status predicted #label MAT\
  146-168      #domain transmembrane #label TM1\
  176-195      #domain transmembrane #label TM2\
  218-241      #domain transmembrane #label TM3\
  256-277      #domain transmembrane #label TM4\
  295-318      #domain transmembrane #label TM5\
  344-363      #domain transmembrane #label TM6\
  376-395      #domain transmembrane #label TM7\
  58,69,100,292 #binding_site carbohydrate (Asn) (covalent) #status
               predicted
SUMMARY       #length 459 #molecular-weight 52057 #checksum 2598
SEQUENCE

```

```

Initial Score   =      5  Optimized Score =      5  Significance =  4.19
Residue Identity = 100%  Matches           =      5  Mismatches   =      0
Gaps            =      0  Conservative Substitutions =      0

```

```

               X  X
               QAEVS
               ||||
RKWRRWHLQGVLCWSSKSQHPWGGSGNGATCSTQVSMLTRVSPSARRSSSFQAEVSLV
      410      420      430      440      450 X  X

```

12. US-08-249-182-4 (1-5)

JN0604 vasoactive intestinal peptide receptor - human

```

ENTRY         JN0604      #type complete
TITLE         vasoactive intestinal peptide receptor - human
ORGANISM      #formal_name Homo sapiens #common_name man
DATE          31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
               31-Dec-1993
ACCESSIONS    JN0604
REFERENCE     JN0604
  #authors     Sreedharan, S.P.; Patel, D.R.; Huang, J.X.; Goetzl, E.J.
  #journal     Biochem. Biophys. Res. Commun. (1993) 193:546-553
  #title       Cloning and functional expression of a human neuroendocrine
               vasoactive intestinal peptide receptor.
  #accession   JN0604
               ##molecule_type mRNA
               ##residues      1-457 ##label SRE
               ##cross-references GB:L13288
               ##note          the nucleotide sequence is not given in this paper
KEYWORDS      glycoprotein; hormone receptor; membrane protein
FEATURE
  135-157      #domain transmembrane #label TM1\
  175-194      #domain transmembrane #label TM2\
  216-240      #domain transmembrane #label TM3\
  255-276      #domain transmembrane #label TM4\
  293-316      #domain transmembrane #label TM5\
  342-361      #domain transmembrane #label TM6\
  374-393      #domain transmembrane #label TM7\

```

38,87,100,270 #binding_site carbohydrate (Asn) (covalent) #status
 predicted
 SUMMARY #length 457 #molecular-weight 51547 #checksum 9283
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 RKWRRWHLQGVLGWNPKYRHPSGGSNGATCSTQVSMLTRVSPGARRSSSFQAEVSLV
 410 420 430 440 450 X

13. US-08-249-182-4 (1-5)

S16562 nolF protein - Rhizobium meliloti

ENTRY S16562 #type complete
 TITLE nolF protein - Rhizobium meliloti
 ORGANISM #formal_name Rhizobium meliloti
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change
 21-Nov-1993
 ACCESSIONS S16562
 REFERENCE S16561
 #authors Baev, N.; Endre, G.; Petrovics, G.; Banfalvi, Z.; Kondorosi,
 A.
 #journal Mol. Gen. Genet. (1991) 228:113-124
 #title Six nodulation genes of nod box locus 4 in Rhizobium meliloti
 are involved in nodulation signal production; nodM codes
 for D-glucosamine synthetase.
 #cross-references MUID:91360053
 #accession S16562
 ##status preliminary
 ##residues 1-367 ##label BAE
 ##cross-references EMBL:X58632
 SUMMARY #length 367 #molecular-weight 39541 #checksum 9824
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.19
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 RGGCLSAQTELAEAVLERNTRLGERGAASEATRLAALADVLDLRAHVRSKQAEVSDAERSLSHAEVRAEFGG
 120 130 140 150 160 X 170 180
 VIRARSVEEGQTVPLNTQLMTIVELNRLEVDAG
 190 200 210

14. US-08-249-182-4 (1-5)

A48470 elongation factor 1 alpha, EF-1 alpha - Eimeria bo

ENTRY A48470 #type fragment
 TITLE elongation factor 1 alpha, EF-1 alpha - Eimeria bovis
 (fragment)
 ORGANISM #formal_name Eimeria bovis
 DATE 01-Dec-1993; #sequence_revision 01-Dec-1993; #text_change
 01-Dec-1993
 ACCESSIONS A48470
 REFERENCE A48470
 #authors Abrahamsen, M.S.; Clark, T.G.; Mascolo, P.; Speer, C.A.;

```

#journal      Mol. Biochem. Parasitol. (1993) 57:1-14
#title        Developmental gene expression in Eimeria bovis.
#cross-references MUID:93149194
#contents      merozoites
#accession     A48470
  ##status      preliminary
  ##molecule_type nucleic acid
  ##residues     1-346 ##label ABR
  ##cross-references NCBIN:123619; NCBIP:123622
  ##note         sequence extracted from NCBI backbone
SUMMARY        #length 346 #checksum 2541
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.19
Residue Identity   =    100%  Matches         =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

                                X  X
                                QAEVS
                                ||||
GFEGAFSKEGQTRHALLAFTLGVKQMIIVGINKMDATTPDKYSETRFNEIQAEVSRYLKTVGYNPEKVPFVP
 20          30          40          50          60          70 X          80

ISGFMGDNMVERSSNMPWYKGIILVEALDNVEP
 90          100          110          120

```

15. US-08-249-182-4 (1-5)

A27659 cytochrome P450 17 - rat (fragment)

```

ENTRY           A27659      #type fragment
TITLE           cytochrome P450 17 - rat (fragment)
ALTERNATE_NAMES cytochrome P450-17-alpha
ORGANISM        #formal_name Rattus norvegicus #common_name Norway rat
DATE            31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change
                18-Jun-1993
ACCESSIONS      A27659
REFERENCE       A27659
  #authors      Nishihara, M.; Winters, C.A.; Buzko, E.; Waternan, M.R.;
                Dufau, M.L.
  #journal      Biochem. Biophys. Res. Commun. (1988) 154:151-158
  #title        Hormonal regulation of rat Leydig cell cytochrome
                P-450-17-alpha mRNA levels and characterization of a
                partial length rat P-450-17-alpha cDNA.
  #cross-references MUID:88280759
  #accession     A27659
    ##molecule_type mRNA
    ##residues     1-237 ##label NIS
    ##note         the authors translated the codon GAT for residues 18,
                131, and 208 as Glu, Asn, and Asn respectively
CLASSIFICATION  #superfamily cytochrome P450
KEYWORDS        heme; monooxygenase; oxidoreductase
FEATURE
  171           #binding_site heme iron (Cys) (axial ligand) #status
                predicted
SUMMARY         #length 237 #checksum 8451
SEQUENCE

```

```

Initial Score      =      5  Optimized Score =      5  Significance =  4.19
Residue Identity   =    100%  Matches         =      5  Mismatches   =      0
Gaps               =      0  Conservative Substitutions =      0

```

```

X  X
QAEVS
||||

```

190 200 210 220 230 X X

> O <
O| |O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_4s.res made by on Thu 22 Sep 94 10:18:01-PDT.

Query sequence being compared:US-08-249-182-4 (1-5)
Number of sequences searched: 36000
Number of scores above cutoff: 4070

Results of the initial comparison of US-08-249-182-4 (1-5) with:
Data bank : Swiss-Prot 28, all entries

100000-
-
N -
U50000-
M -
B -
E -
R -
-
O -
F10000-
-
S -
E 5000*
O -
U -
E -
N -
C -
E -
S 1000-
-
-
500-
-
-
-
-
-
100-
-
-
50-
-
-
-
-
10-
-
-
5-
-
-
-

*

*

*

*

*


```

-
-
0-----
|| | | | | | | | | |
SCORE 0| 1 |1 2 | 2 3 | 3 4 4 5
STDEV 0 1 2 4

```

PARAMETERS

```

Similarity matrix      Unitary      K-tuple      2
Mismatch penalty      1      Joining penalty      20
Gap penalty            1.00      Window size      5
Gap size penalty      0.05
Cutoff score          0
Randomization group    0

Initial scores to save      40      Alignments to save      15
Optimized scores to save    0      Display context      50

```

SEARCH STATISTICS

```

Scores:      Mean      Median      Standard Deviation
              1          3          0.84

```

```

Times:      CPU      Total Elapsed
            00:00:49.92      00:00:54.00

```

```

Number of residues:      12496420
Number of sequences searched:      36000
Number of scores above cutoff:      4070

```

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

8 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. ENPL_CATRO	ENDOPLASMIN HOMOLOG PRECURSOR	817	5	5	4.75	0
2. RAL2_SCHPD	RAL2 PROTEIN.	611	5	5	4.75	0
3. IBMP_FMVD	INCLUSION BODY MATRIX PROTEIN	512	5	5	4.75	0
4. CPT7_RAT	CYTOCHROME P450 XVIIIA1 (P450-	507	5	5	4.75	0
5. VIPR_RAT	VASOACTIVE INTESTINAL POLYPEP	459	5	5	4.75	0
6. VIPR_HUMAN	VASOACTIVE INTESTINAL POLYPEP	457	5	5	4.75	0
7. NOLF_RHME	NODULATION PROTEIN NOLF.	367	5	5	4.75	0
8. YM2_STRCD	MINI-CIRCLE HYPOTHETICAL 13.3	122	5	5	4.75	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 3 standard deviations above mean ****						
9. RP30_YEAST	RIBOSOMAL PROTEIN RP30 (FRAGM	25	4	4	3.56	0
10. TYBB_ONCMY	THYMOSIN BETA-12.	42	4	4	3.56	0
11. LCA_MACGI	ALPHA-LACTALBUMIN I (LACTOSE	42	4	4	3.56	0
12. THIO_EUBAC	THIOREDOXIN (FRAGMENT).	45	4	4	3.56	0
13. PSBK_CHLRE	PHOTOSYSTEM II 4 KD REACTION	46	4	4	3.56	0

15.	VG18_BPML5	GENE 18 PROTEIN (GP18).	57	4	4	3.56	0
16.	PSBK_TOBAC	PHOTOSYSTEM II 4 KD REACTION	61	4	4	3.56	0
17.	PSBK_SINAL	PHOTOSYSTEM II 4 KD REACTION	61	4	4	3.56	0
18.	PSBK_ORYSA	PHOTOSYSTEM II 4 KD REACTION	61	4	4	3.56	0
19.	PSBK_HORVU	PHOTOSYSTEM II 4 KD REACTION	61	4	4	3.56	0
20.	RINB_BPPHA	TRANSCRIPTIONAL ACTIVATOR RIN	62	4	4	3.56	0
21.	HST_YEREN	HEAT-STABLE ENTEROTOXIN PRECU	71	4	4	3.56	0
22.	CPB2_ECOLI	COPB PROTEIN (REPA2 PROTEIN).	84	4	4	3.56	0
23.	YPB2_ECOLI	HYPOTHETICAL 10.0 KD PROTEIN.	87	4	4	3.56	0
24.	Y622_BPT4	HYPOTHETICAL 10.9 KD PROTEIN	92	4	4	3.56	0
25.	E111_ADEM1	EARLY E1A 11 KD PROTEIN.	96	4	4	3.56	0
26.	HG14_HUMAN	NONHISTONE CHROMOSOMAL PROTEI	99	4	4	3.56	0
27.	HG14_BOVIN	NONHISTONE CHROMOSOMAL PROTEI	100	4	4	3.56	0
28.	YKE4_YEAST	HYPOTHETICAL 11.4 KD PROTEIN	101	4	4	3.56	0
29.	PT1_BACSU	PHOSPHOENOLPYRUVATE-PROTEIN P	102	4	4	3.56	0
30.	CCMK_SYNP7	CARBON DIOXIDE CONCENTRATING	102	4	4	3.56	0
31.	THI1_YEAST	THIOREDOXIN I (TR-I).	104	4	4	3.56	0
32.	PUTX_PSEPU	PUTIDAREDOXIN (PDX).	106	4	4	3.56	0
33.	INS1_XENLA	INSULIN 1 PRECURSOR.	106	4	4	3.56	0
34.	YLC6_YEREN	HYPOTHETICAL 12.2 KD PROTEIN	110	4	4	3.56	0
35.	INS_CAVPO	INSULIN PRECURSOR.	110	4	4	3.56	0
36.	GVPK_HALHA	GVPK PROTEIN.	113	4	4	3.56	0
37.	Y13K_TYLCM	HYPOTHETICAL 13.3 KD PROTEIN	115	4	4	3.56	0
38.	HV3N_HUMAN	IG HEAVY CHAIN V-III REGION (119	4	4	3.56	0
39.	LCA_MACRG	ALPHA-LACTALBUMIN (LACTOSE SY	121	4	4	3.56	0
40.	TRM2_ECOLI	TRAM PROTEIN.	127	4	4	3.56	0

1. US-08-249-182-4 (1-5)

ENPL_CATRO ENDOPLASMIN HOMOLOG PRECURSOR (HEAT SHOCK 90 KD PR

ID ENPL_CATRO STANDARD; PRT; 817 AA.
AC P35016;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE ENDOPLASMIN HOMOLOG PRECURSOR (HEAT SHOCK 90 KD PROTEIN).
GN HSP90.
OS CATHARANTHUS ROSEUS (ROSY PERIWINKLE) (MADAGASCAR PERIWINKLE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC GENTIANALES; APOCYNACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CP3A;
RM 94033337
RA SCHROEDER G., BECK M., EICHEL J., VETTER H.P., SCHROEDER J.;
RL PLANT MOL. BIOL. 23:583-594(1993).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN HSP90 FAMILY.
DR EMBL; L14594; CTRHSP90A.
KW CHAPERONE; ENDOPLASMIC RETICULUM; HEAT SHOCK; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 21 795 ENDOPLASMIN HOMOLOG.
FT CARBOHYD 111 111 POTENTIAL.
FT CARBOHYD 306 306 POTENTIAL.
FT CARBOHYD 416 416 POTENTIAL.
FT CARBOHYD 456 456 POTENTIAL.
FT CARBOHYD 624 624 POTENTIAL.
FT SITE 814 817 PREVENT SECRETION FROM ER.
SQ SEQUENCE 817 AA; 93491 MW; 3430345 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

A A
0AEVS
|||||

DSDAPVDPKVEDKIGAVPNGLSTDSDVAKREAESMSMRNLRSDAEKFEFGAEVSRMLMDIINSLYSNKDIF
 40 50 60 70 80 X X 90 100

LRELISNASDALDKIRFLALTDKEILGEGDTAK
 110 120 130

2. US-08-249-182-4 (1-5)

RAL2_SCHPD RAL2 PROTEIN.

ID RAL2_SCHPD STANDARD; PRT; 611 AA.
 AC P15258;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE RAL2 PROTEIN.
 GN RAL2.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90066514
 RA FUKUI Y., MIYAKE S., SATOH M., YAMAMOTO M.;
 RL MOL. CELL. BIOL. 9:5617-5622(1989).
 CC -!- FUNCTION: IMPLICATED IN ACTIVATION OF THE RAS1 PROTEIN. IT IS
 CC PROBABLY A GDP-GTP EXCHANGE PROTEIN FOR RAS1.
 DR EMBL; M30827; SPAL2.
 DR PIR; A33827; A33827.
 KW GUANINE-NUCLEOTIDE RELEASING FACTOR.
 SQ SEQUENCE 611 AA; 69847 MW; 1910377 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
0AEVS
|||||

GKLLLNGLISDMEILTIERMHIPCLSRMLYKRWPAFQKIMDRAVEKNQEAFAEVSQGLGPQLTDLPFSSIHST
 430 440 450 460 470 480 490

GSRALYMPYSFETCSAFLHYIYCGTLNGSYCTA
 500 510 520 530

3. US-08-249-182-4 (1-5)

IBMP_FMVD INCLUSION BODY MATRIX PROTEIN (VIROPLASMIN).

ID IBMP_FMVD STANDARD; PRT; 512 AA.
 AC P09524;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-AUG-1991 (REL. 19, LAST ANNOTATION UPDATE)
 DE INCLUSION BODY MATRIX PROTEIN (VIROPLASMIN).
 GN VI.
 OS FIGWORT MOSAIC VIRUS (STRAIN DXS) (FMV).
 OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; CAULIMOVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 88040466
 RA RICHINS R.D., SCHOLTHOF H.B., SHEPHERD R.J.;
 RL NUCLEIC ACIDS RES. 15:8451-8466(1987).
 CC -!- FUNCTION: ENHANCES THE TRANSLATION OF DOWNSTREAM ORF'S ON

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC INCLUSION BODIES.
 CC -!- THE INCLUSION BODIES ARE THE SITE OF VIRAL DNA SYNTHESIS, VIRION
 CC ASSEMBLY AND ACCUMULATION IN THE INFECTED CELL.
 CC -!- SIMILARITY: HIGH, WITH OTHER CAULIMOVIRUS VIROPLASMIN.
 DR EMBL; X06166; CAFMVXX.
 DR PIR; S01284; WMCVFM.
 KW TRANS-ACTING FACTOR; TRANSLATION REGULATION.
 SQ SEQUENCE 512 AA; 58207 MW; 1378914 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 GAEVS
 |||||

FKPDYLRASNGGSWFAVYKGPKNKEFFTEWEIVADICKKRQKSKRFRSKEGAEVSISLYNKDIQDPVNFLRP
 150 160 170 180 190 X X 200 210

VKLVKEERAAQPLKFKAIAAEQTIQFDEFRQIW
 220 230 240

4. US-08-249-182-4 (1-5)

CPT7_RAT CYTOCHROME P450 XVIIIA1 (P450-C17) (EC 1.14.99.9) (

ID CPT7_RAT STANDARD; PRT; 507 AA.
 AC P11715;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE CYTOCHROME P450 XVIIIA1 (P450-C17) (EC 1.14.99.9) (STEROID 17-ALPHA-
 DE HYDROXYLASE/17,20 LYASE).
 GN CYP17.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 89295447
 RA FEVOLD H.R., LORENCE M.C., MCCARTHY J.L., TRANT J.M., KAGIMOTO M.,
 RA WATERMAN M.R., MASON J.I.;
 RL MOL. ENDOCRINOL. 3:968-975(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RM 89076306
 RA NAMIKI M., KITAMURA M., BUCZKO E., DUFAY M.L.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 157:705-712(1988).
 RN [3]
 RP SEQUENCE OF 271-507 FROM N.A.
 RM 88280759
 RA NISHIHARA M., WINTERS C.A., BUZKO E., WATERMAN M.R., DUFAY M.L.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 154:151-158(1988).
 RN [4]
 RP SEQUENCE OF 273-507 FROM N.A.
 RM 90046678
 RA MELLON S.H., VAISSE C.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:7775-7779(1989).
 CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 CC MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED
 CC COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.
 CC -!- CATALYTIC ACTIVITY: A STEROID + AH(2) + O(2) = A 17-ALPHA-
 CC HYDROXYSTEROID + A + H(2)O.
 DR EMBL; M31681; RNP45017.

DR EMBL; M22204; RNP430A.
 DR EMBL; M27282; RNP450C1.
 DR PIR; A27659; A27659.
 DR PIR; A30828; A30828.
 DR PIR; S16719; S16719.
 DR PROSITE; PS00086; CYTOCHROME_P450.
 KW ELECTRON TRANSPORT; OXIDOREDUCTASE; MONOOXYGENASE; MEMBRANE;
 KW HEME; STEROIDOGENESIS.
 FT BINDING 441 441 HEME.
 SQ SEQUENCE 507 AA; 57250 MW; 1347075 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 GAEVS
 I I I I
 FVFTALLLRFDLDVSDDKQLPRLEGDPKVVFLIDPFVKVITVRGAWMDAGAEVST
 460 470 480 490 500 X X

5. US-08-249-182-4 (1-5)

VIPR_RAT VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECU

ID VIPR_RAT STANDARD; PRT; 459 AA.
 AC P30083;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR (VIP-R-1).
 OS RATTUS NORVEGICUS (RAT).
 DC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RM 92232309
 RA ISHIHARA T., SHIGEMOTO R., MORI K., TAKAHASHI K., NAGATA S.;
 RL NEURON 8:811-819(1992).
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CC CYCLASE.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: IN LIVER, LUNG, INTESTINES, THYMUS AND BRAIN
 CC (MOSTLY IN THE CEREBRAL CORTEX AND HIPPOCAMPUS).
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; M86835; RNVASREC.
 DR PIR; JH0594; JH0594.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 459 VASOACTIVE INTESTINAL POLYPEPTIDE
 FT RECEPTOR 1.
 FT DOMAIN 31 143 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 144 168 1 (POTENTIAL).
 FT DOMAIN 169 175 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 176 195 2 (POTENTIAL).
 FT DOMAIN 196 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 241 3 (POTENTIAL).
 FT DOMAIN 242 255 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 256 277 4 (POTENTIAL).
 FT DOMAIN 278 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 295 318 5 (POTENTIAL).

FT DOMAIN 317 341 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 342 361 6 (POTENTIAL).
 FT DOMAIN 362 373 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 374 393 7 (POTENTIAL).
 FT DOMAIN 394 457 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 58 58 POTENTIAL.
 FT CARBOHYD 69 69 POTENTIAL.
 FT CARBOHYD 100 100 POTENTIAL.
 FT CARBOHYD 290 290 POTENTIAL.
 SQ SEQUENCE 457 AA; 51547 MW; 1161417 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 RKWRRWHLQGVLGWNPKYRHPSGGSNGATCSTQVSMLTRVSPGARRSSSFQAEVSLV
 410 420 430 440 450 X

7. US-08-249-182-4 (1-5)

NOLF_RHIME NODULATION PROTEIN NOLF.

ID NOLF_RHIME STANDARD; PRT; 367 AA.
 AC P25196;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE NODULATION PROTEIN NOLF.
 GN NOLF.
 OS RHIZOBIUM MELILOTI.
 OG PLASMID SYM PRME41B.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC RHIZOBIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AK 631;
 RM 91360053
 RA BAEV N., ENDRE G., PETROVICS G., BANFALVI Z., KONDOROSI A.;
 RL MOL. GEN. GENET. 228:113-124(1991).
 CC -!- FUNCTION: INVOLVED IN THE PRODUCTION OF MEDICAGO-SPECIFIC
 CC NODULATION SIGNAL MOLECULE.
 DR EMBL; X58632; RMPRME41B.
 DR PIR; S16562; S16562.
 KW PLASMID; NODULATION.
 SQ SEQUENCE 367 AA; 39541 MW; 661978 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.75
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 ||||
 RGGCLSAQTELAEAVLERNTRLGERGAASEATRLAALADVLDLRAHVRSKQAEVSDAERSLSHAEVRAEFEGG
 120 130 140 150 160 X 170 180
 VIRARSVEEGQTVPLNTQLMTIVELNRLEVDAG
 190 200 210

8. US-08-249-182-4 (1-5)

YM2_STRCO MINI-CIRCLE HYPOTHETICAL 13.3 KD PROTEIN.

FT	DOMAIN	317	343	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	344	363	6 (POTENTIAL).
FT	DOMAIN	364	375	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	376	395	7 (POTENTIAL).
FT	DOMAIN	396	459	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	58	58	POTENTIAL.
FT	CARBOHYD	69	69	POTENTIAL.
FT	CARBOHYD	100	100	POTENTIAL.
FT	CARBOHYD	292	292	POTENTIAL.
SO	SEQUENCE	459 AA; 52057 MW; 1186786 CN;		

Initial Score = 5 Optimized Score = 5 Significance = 4.75
Residue Identity = 100% Matches = 5 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

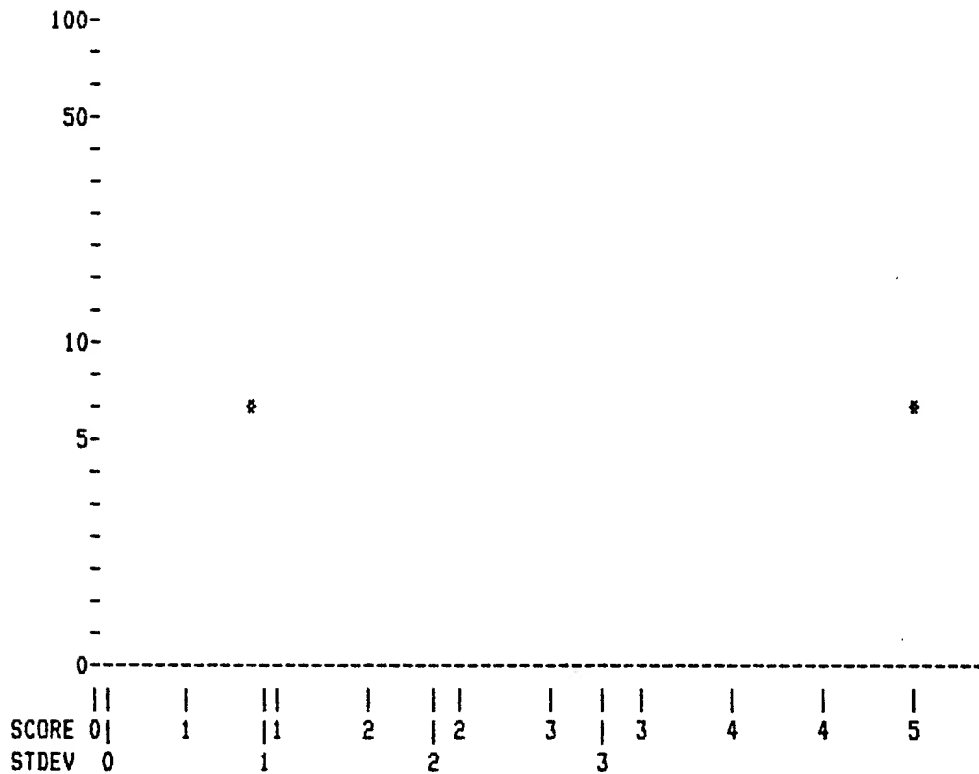
                                X  X
                                GAEVS
                                |||||
      RKWRRWHLQGVLGWSSKSGHPWGGSNATCSTQVSNLTVSPSARRSSSFQAEVSLV
      410          420          430          440          450 X  X

```

6. US-08-249-182-4 (1-5)

VIPR_HUMAN VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECU

ID VIPR_HUMAN STANDARD; PRT; 457 AA.
AC P32241;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR (VIP-R-1).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RM 93290641
RA SREEDHARAN S.P., PATEL D.R., HUANG J.-X., GOETZL E.J.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 193:546-553(1993).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
CC CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: IN LUNG, HT29 COLONIC EPITHELIAL CELLS,
CC RAJI B-LYMPHOBLASTS. LESSER EXTENT IN BRAIN, HEART, KIDNEY,
CC LIVER AND PLACENTA.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; L13288; HSVIPR1.
DR PIR; JN0604; JN0604.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 457 VASOACTIVE INTESTINAL POLYPEPTIDE
FT RECEPTOR 1.
FT DOMAIN 31 142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 143 167 1 (POTENTIAL).
FT DOMAIN 168 174 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 175 194 2 (POTENTIAL).
FT DOMAIN 195 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 240 3 (POTENTIAL).
FT DOMAIN 241 254 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 255 276 4 (POTENTIAL).
FT DOMAIN 277 292 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 293 316 5 (POTENTIAL).



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.88

Times:	CPU	Total Elapsed
	00:00:48.91	00:00:52.00

Number of residues:	12496420
Number of sequences searched:	36000
Number of scores above cutoff:	3832

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

6 100% similar sequences to the query sequence were found:

Init. Opt.

Sequence Name	Description	Length	Score	Score	Sig.	Frame
1. PULA_KLEAE	PULLULANASE (EC 3.2.1.41) (AL	1096	5	5	4.52	0
2. PRIS_DESVH	PRISMANE PROTEIN.	553	5	5	4.52	0
3. PRIS_DESDE	PRISMANE PROTEIN.	544	5	5	4.52	0
4. CPXL_PSESP	CYTOCHROME P450-TERP (EC 1.14	428	5	5	4.52	0
5. YBSB_YEAST	HYPOTHETICAL 47.4 KD PROTEIN	418	5	5	4.52	0
6. DBH_THETH	DNA-BINDING PROTEIN II.	95	5	5	4.52	0

The list of other best scores is:

Init. Opt.

Sequence Name	Description	Length	Score	Score	Sig.	Frame
**** 3 standard deviations above mean ****						
7. C551_PSEST	CYTOCHROME C551.	82	4	4	3.39	0
8. CYC6_PLEBO	CYTOCHROME C6 (SOLUBLE CYTOCH	85	4	4	3.39	0
9. CYC6_ANAVA	CYTOCHROME C6 (SOLUBLE CYTOCH	86	4	4	3.39	0
10. DBH_BACST	DNA-BINDING PROTEIN II (HB) (90	4	4	3.39	0
11. DBH_BACSU	DNA-BINDING PROTEIN II (HB) (92	4	4	3.39	0
12. PINO_ECOLI	PINO PROTEIN.	105	4	4	3.39	0
13. GLN1_METTL	NITROGEN FIXATION NIFHD REGIO	105	4	4	3.39	0
14. YP12_RTBPV	HYPOTHETICAL P12 PROTEIN (DRF	110	4	4	3.39	0
15. YP12_RTBV	HYPOTHETICAL P12 PROTEIN (DRF	110	4	4	3.39	0
16. CYC6_ANASQ	CYTOCHROME C6 PRECURSOR (SOLU	111	4	4	3.39	0
17. CYC6_ANASP	CYTOCHROME C6 PRECURSOR (SOLU	111	4	4	3.39	0
18. ND75_PEA	EARLY NODULIN-75 PROTEIN (N-7	112	4	4	3.39	0
19. YCW8_YEAST	HYPOTHETICAL 12.4 KD PROTEIN	114	4	4	3.39	0
20. WNT3_EPTST	WNT-3 PROTEIN (FRAGMENT).	123	4	4	3.39	0
21. CY2_RHOCA	CYTOCHROME C2 PRECURSOR.	137	4	4	3.39	0
22. YZP2_ECOLI	VERY HYPOTHETICAL 16.1 KD PRO	146	4	4	3.39	0
23. PETD_PROHO	CYTOCHROME B6-F COMPLEX SUBUN	160	4	4	3.39	0
24. MAT1_YEAST	MATING-TYPE PROTEIN ALPHA-1.	175	4	4	3.39	0
25. SPC4_CANFA	MICROSOMAL SIGNAL PEPTIDASE 1	179	4	4	3.39	0
26. SPC3_CANFA	MICROSOMAL SIGNAL PEPTIDASE 2	191	4	4	3.39	0
27. COAG_LIMPO	COAGULOGEN PRECURSOR (CONTAIN	195	4	4	3.39	0
28. RL22_SPIOL	50S RIBOSOMAL PROTEIN L22 (RI	199	4	4	3.39	0
29. COAT_BYDVN	COAT PROTEIN.	204	4	4	3.39	0
30. COAT_PLRVW	COAT PROTEIN.	208	4	4	3.39	0
31. COAT_PLRVR	COAT PROTEIN.	208	4	4	3.39	0
32. COAT_PLRV1	COAT PROTEIN.	208	4	4	3.39	0
33. COAT_PLRV	COAT PROTEIN.	208	4	4	3.39	0
34. HUPD_RHOCA	HUPD PROTEIN.	210	4	4	3.39	0
35. CAT1_STAAV	CHLORAMPHENICOL ACETYLTRANSFE	216	4	4	3.39	0
36. CAT_ECOLI	CHLORAMPHENICOL ACETYLTRANSFE	219	4	4	3.39	0
37. UL01_HSV11	GLYCOPROTEIN L PRECURSOR.	224	4	4	3.39	0
38. FRDB_WOLSV	FUMARATE REDUCTASE IRON-SULFU	239	4	4	3.39	0
39. DNAC_ECOLI	DNAC PROTEIN.	245	4	4	3.39	0
40. NU6M_WHEAT	NADH-UBIQUINONE OXIDOREDUCTAS	247	4	4	3.39	0

1. US-08-249-182-3 (1-5)

PULA_KLEAE PULLULANASE (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-

ID PULA_KLEAE STANDARD; PRT: 1096 AA.
AC P07811;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE PULLULANASE (EC 3.2.1.41) (ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)
DE (PULLULAN 6-GLUCANOHYDROLASE).
GN PULA.
OS KLEBSIELLA AEROGENES.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W70;
 RM 87194626
 RA KATSURAGI N., TAKIZAWA N., MUROOKA Y.;
 RL J. BACTERIOL. 169:2301-2306(1987).
 CC -!- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYZES
 CC (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO
 CC FORM MALTOTRIOSE.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (PROBABLE).
 CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL; M16187; KAPULA.
 DR PIR; A26879; A26879.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN.
 KW HYDROLASE; GLYCOSIDASE; MEMBRANE; LIPOPROTEIN; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 1096 PULLULANASE.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE.
 FT ACT_SITE 694 694 BY SIMILARITY.
 FT ACT_SITE 723 723 BY SIMILARITY.
 FT ACT_SITE 851 851 BY SIMILARITY.
 SQ SEQUENCE 1096 AA; 119335 MW; 5852107 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||||
 GENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLOGDDGG
 210 220 230 240 250 260 X 270 280
 IAESDGILSLSHPGADRRRAGRYLCRRAEALSY
 290 300 310

2. US-08-249-182-3 (1-5)

PRIS_DESVH PRISMANE PROTEIN.

ID PRIS_DESVH STANDARD; PRT; 553 AA.
 AC P31101;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PRISMANE PROTEIN.
 OS DESULFOVIBRIO VULGARIS (STRAIN HILDENBOROUGH).
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
 OC SULFATE- OR SULFUR-REDUCING DISSIMILATORY.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RM 92394141
 RA STOKKERMANS J.P.W.G., PIERIK A.J., WOLBERT R.B.G., HAGEN W.R.,
 RA VAN DONGEN W.M.A.M., VEEGER C.;
 RL EUR. J. BIOCHEM. 208:435-442(1992).
 RN [2]
 RP CHARACTERIZATION, AND SEQUENCE OF 1-14.
 RC STRAIN=NCIB 8303;
 RM 92298997
 RA PIERIK A.J., WOLBERT R.B.G., MUTSAERS P.H.A., HAGEN W.R., VEEGER C.;
 RL EUR. J. BIOCHEM. 206:697-704(1992).
 RN [3]
 RP CHARACTERIZATION.

RC STRAIN=NCIB 8303;
 RM 92298998
 RA PIERIK A.J., HAGEN W.R., DUNHAM W.R., SANDS R.H.;
 RL EUR. J. BIOCHEM. 206:705-719(1992).
 CC -!- FUNCTION: PROBABLE ENZYME OF YET UNCHARACTERIZED ACTIVITY.
 CC -!- COFACTOR: CONTAINS A [6FE-6S] PRISMANE CLUSTER THAT MIGHT OCCUR
 CC IN 4 DIFFERENT REDOX STATES: [6FE-6S]3+; [6FE-6S]4+; [6FE-6S]5+
 CC AND [6FE-6S]6+.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 DR EMBL; Z11707; DVPRIS.
 KW ELECTRON TRANSPORT; IRON-SULFUR.
 FT METAL 3 3 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 6 6 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 15 15 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 21 21 IRON-SULFUR (6FE-6S) (POTENTIAL).
 SQ SEQUENCE 553 AA; 60163 MW; 1555801 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 |||||
 EITQVNIGVGKNPGILISGHDLDKDMAELLKQTEGTGVDVYTHGEHLPANYYPAFKKYPHFVGNVGGSWWQON
 230 240 250 260 270 X 280 290
 PEFESFNGPILLTTNCLVPLKKENTYLDRLYTT
 300 310 320

3. US-08-249-182-3 (1-5)

PRIS_DESDE PRISMANE PROTEIN.

ID PRIS_DESDE STANDARD; PRT; 544 AA.
 AC Q01770;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PRISMANE PROTEIN.
 OS DESULFOVIBRIO DESULFURICANS.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
 OC SULFATE- OR SULFUR-REDUCING DISSIMILATORY.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27774;
 RM 92379097
 RA STOKKERMANS J.P.W.G., VAN DEN BERG W.A.M., VAN DONGEN W.M.A.M.,
 RA VEEGER C.;
 RL BIOCHIM. BIOPHYS. ACTA 1132:83-87(1992).
 RN [2]
 RP PARTIAL SEQUENCE OF 1-36, AND CHARACTERIZATION.
 RC STRAIN=ATCC 27774;
 RM 92165800
 RA MOURA I., TAVARES P., MOURA J.J.G., RAVI N., HUYNH B.H., LIU M.-Y.,
 RA LE GALL J.;
 RL J. BIOL. CHEM. 267:4489-4496(1992).
 CC -!- FUNCTION: PROBABLE ENZYME OF YET UNCHARACTERIZED ACTIVITY.
 CC -!- COFACTOR: CONTAINS A [6FE-6S] PRISMANE CLUSTER THAT MIGHT OCCUR
 CC IN 4 DIFFERENT REDOX STATES: [6FE-6S]3+; [6FE-6S]4+; [6FE-6S]5+
 CC AND [6FE-6S]6+.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 DR EMBL; Z11975; DSPRISMAN.
 DR PIR; A42396; A42396.

KW ELECTRON TRANSPORT; IRON-SULFUR.
 FT INIT_MET 0 0
 FT METAL 6 6 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 9 9 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 18 18 IRON-SULFUR (6FE-6S) (POTENTIAL).
 FT METAL 24 24 IRON-SULFUR (6FE-6S) (POTENTIAL).
 SQ SEQUENCE 544 AA; 58528 MW; 1518483 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 EITKVNIGVGSNPGILISGHLRLDLEMLLKQTEGTGVDVYTHSEMLPAHYYPAFKKYAHFKGNYGNANWKQK
 230 240 250 260 270 X 280 290
 EEFESFNGPVLTTNCLVPPKDSYKDRVYTTGI
 300 310 320

4. US-08-249-182-3 (1-5)
 CPXL_PSESP CYTOCHROME P450-TERP (EC 1.14.-.-).

ID CPXL_PSESP STANDARD; PRT; 428 AA.
 AC P33006;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE CYTOCHROME P450-TERP (EC 1.14.-.-).
 GN CYP108 OR TERPC.
 OS PSEUDOMONAS SP.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC PSEUDOMONADACEAE.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RM 92332528
 RA PETERSON J.A., LU J.-Y., GEISSELSODER J., GRAHAM-LORENCE S.,
 RA CARMONA C., WITNEY F., LORENCE M.C.;
 RL J. BIOL. CHEM. 267:14193-14203(1992).
 CC -!- FUNCTION: INVOLVED IN A ALPHA-TERPINEOL OXIDATION SYSTEM.
 CC -!- SIMILARITY: MEMBER OF THE CYTOCHROME P-450 FAMILY.
 DR EMBL; M91440; PSTERP.
 DR EMBL; S39894; S39894.
 DR PIR; S27653; S27653.
 DR PROSITE; PS00086; CYTOCHROME_P450.
 KW OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME.
 FT BINDING 377 377 HEME (BY SIMILARITY).
 SQ SEQUENCE 428 AA; 47922 MW; 907644 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 MDARATPEHIARTVILPQGYADDEVIYPAFKWLRLDEQPLAMAHIEGYDPMWIATKHADVMDIGKQPGLFNS
 10 20 30 X 40 50 60 70
 AEGSEILYDQ
 80

5. US-08-249-182-3 (1-5)

YB5B_YEAST HYPOTHETICAL 47.4 KD PROTEIN IN KIP1-SEC17 INTERGE

ID YB5B_YEAST STANDARD; PRT; 418 AA.
 AC P34220;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 47.4 KD PROTEIN IN KIP1-SEC17 INTERGENIC REGION.
 GN YBL0511.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA SCHERENS B., EL BAKKOURY M., VIERENDEELS F., DUBOIS E., MESSENGUY F.;
 RL YEAST 9:1355-1371(1993).
 DR EMBL; 723261; SCIL260.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 418 AA; 47390 MW; 911132 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||

SLRTEENLAVVKQIPTEKLLLETDAPWCEIKRTHASFQYLAKYQEVDFEYPAFKSVKKNKLADKLNAEELY
 310 320 330 340 350 X 360 370

MVKGRNEPCNMEQVAIVVSEVKDVLATLIDTT
 380 390 400

6. US-08-249-182-3 (1-5)

DBH_THETH DNA-BINDING PROTEIN II.

ID DBH_THETH STANDARD; PRT; 95 AA.
 AC P19436;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
 DE DNA-BINDING PROTEIN II.
 OS THERMUS AQUATICUS (SUBSP. THERMOPHILUS).
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC UNCERTAIN.
 RN [1]
 RP SEQUENCE.
 RM 91032203
 RA ZIERER R., CHOLI D.;
 RL FEBS LETT. 273:59-62(1990).
 CC -!- FUNCTION: THIS PROTEIN BELONGS OT THE HISTONE LIKE FAMILY OF
 CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING
 CC DNA AND TO STABILIZE DNA FROM DENATURATION UNDER EXTREME
 CC ENVIRONMENTAL CONDITIONS.
 CC -!- SUBUNIT: HOMODIMER.
 DR PIR; S12888; S12888.
 DR PROSITE; PS00045; HISTONE_LIKE.
 KW DNA-BINDING; DNA CONDENSATION.
 SQ SEQUENCE 95 AA; 10163 MW; 48066 CN;

Initial Score = 5 Optimized Score = 5 Significance = 4.52
 Residue Identity = 100% Matches = 5 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

A A
YPAFK
|||||

DALLAKVEEALANGSKVQLTGFGTFEVRKRRKARTGVKPGTKEKIKIPATQYPAFKPGKALKDKVK

40 50 60 70 80 X 90

7. US-08-249-182-3 (1-5)

C551_PSEST CYTOCHROME C551.

ID C551_PSEST STANDARD; PRT; 82 AA.
AC P00101;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE CYTOCHROME C551.
GN NIRM.
OS PSEUDOMONAS STUTZERI (PSEUDOMONAS PERFECTOMARINA).
DC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
DC PSEUDOMONADACEAE.
RN [1]
RP SEQUENCE.
RC STRAIN=221;
RM 73224976
RA AMBLER R.P., WYNN M.;
RL BIOCHEM. J. 131:485-498(1973).
RN [2]
RP STRUCTURE BY NMR.
RC STRAIN=ATCC 17588;
RM 93002990
RA CAI M., BRADFORD E.G., TIMKOVICH R.;
RL BIOCHEMISTRY 31:8603-8612(1992).
CC -!- FUNCTION: ELECTRON DONOR FOR CYTOCHROME CD1 IN NITRITE AND NITRATE
CC RESPIRATION.
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -!- SIMILARITY: 11 DIFFERENCES WITH STRAIN ZOBELL.
DR PIR; A00093; CCPS55.
DR PROSITE; PS00190; CYTOCHROME_C.
KW ELECTRON TRANSPORT; OXIDATIVE PHOSPHORYLATION; HEME; PERIPLASMIC.
FT BINDING 12 12 HEME (COVALENT).
FT BINDING 15 15 HEME (COVALENT).
FT METAL 16 16 IRON (HEME AXIAL LIGAND).
FT METAL 61 61 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 82 AA; 8612 MW; 33580 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
YPAFK
|||||

QDGEALFKSKPCAACHSIDAKLVGPAFKEVAAKYAGQDGAADLLAGHIKNGSQGVWGPIPMPPNPVTEEEAK

10 20 X 30 40 50 60 70

ILAEWI

8. US-08-249-182-3 (1-5)

CYC6_PLEBD CYTOCHROME C6 (SOLUBLE CYTOCHROME F) (CYTOCHROME C

ID CYC6_PLEBD STANDARD; PRT; 85 AA.
AC P00117;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)

DT 01-MAR-1989 (REL. 10, LAST ANNOTATION UPDATE)
 DE CYTOCHROME C6 (SOLUBLE CYTOCHROME F) (CYTOCHROME C553).
 OS PLECTONEMA BORYANUM.
 OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
 DC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CCAP 1462/2;
 RM 78023897
 RA AITKEN A.;
 RL EUR. J. BIOCHEM. 78:273-279(1977).
 CC -!- FUNCTION: CYTOCHROME C6 IS A MONOHEME MONOMER. IT FUNCTIONS AS AN
 CC ELECTRON CARRIER BETWEEN MEMBRANE-BOUND CYTOCHROME F AND P700 IN
 CC THE PHOTOPHOSPHORYLATION CHAIN IN CHLOROPLASTS AND ALGAE. IT
 CC SUBSTITUTES FOR PLASTOCYANIN IN COPPER-DEFICIENT BLUE-GREEN ALGAE
 CC AND IN THE CHLOROPLASTS OF SOME EUKARYOTE ALGAE.
 DR PIR; A00109; CCPB6.
 DR PROSITE; PS00190; CYTOCHROME_C.
 KW ELECTRON TRANSPORT; PHOTOSYNTHESIS; HEME.
 FT BINDING 14 14 HEME (COVALENT).
 FT BINDING 17 17 HEME (COVALENT).
 FT METAL 18 18 IRON (HEME AXIAL LIGAND).
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 85 AA; 8576 MW; 34505 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 KVFNANCAACHASGGGQINGAKTLKKNALTANGKDTVEAIVAQVTNGKGAMPAFKGRLSDDQIQSVALYVLD
 10 20 30 40 50 60 X 70
 KAEKGW
 80

9. US-08-249-182-3 (1-5)

CYC6_ANAVA CYTOCHROME C6 (SOLUBLE CYTOCHROME F) (CYTOCHROME C

ID CYC6_ANAVA STANDARD; PRT; 86 AA.
 AC P00113;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-MAR-1989 (REL. 10, LAST ANNOTATION UPDATE)
 DE CYTOCHROME C6 (SOLUBLE CYTOCHROME F) (CYTOCHROME C553).
 OS ANABAENA VARIABILIS.
 OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
 DC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
 RN [1]
 RP SEQUENCE.
 RM 82265568
 RA BEECHER J., MARGOLIAH E.;
 RL UNPUBLISHED RESULTS, CITED BY:
 RL ULRICH E.L., KROGMANN D.W., MARKLEY J.L.;
 RL J. BIOL. CHEM. 257:9356-9364(1982).
 RN [2]
 RP SEQUENCE OF 1-22; 30-39 AND 56-86.
 RM 77056395
 RA AITKEN A.;
 RL NATURE 263:793-796(1976).
 CC -!- FUNCTION: CYTOCHROME C6 IS A MONOHEME MONOMER. IT FUNCTIONS AS AN
 CC ELECTRON CARRIER BETWEEN MEMBRANE-BOUND CYTOCHROME F AND P700 IN
 CC THE PHOTOPHOSPHORYLATION CHAIN IN CHLOROPLASTS AND ALGAE. IT

CC SUBSTITUTES FOR PLASTOCYANIN IN COPPER-DEFICIENT BLUE-GREEN ALGAE
 CC AND IN THE CHLOROPLASTS OF SOME EUKARYOTE ALGAE.
 DR PIR; A00105; CCA16.
 DR PROSITE; PS00190; CYTOCHROME_C.
 KW ELECTRON TRANSPORT; PHOTOSYNTHESIS; HEME.
 FT BINDING 14 14 HEME (COVALENT).
 FT BINDING 17 17 HEME (COVALENT).
 FT METAL 18 18 IRON (HEME AXIAL LIGAND).
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).
 FT CONFLICT 82 82 E -> D (IN REF. 2).
 FT CONFLICT 84 84 E -> D (IN REF. 2).
 SQ SEQUENCE 86 AA; 8973 MW; 36422 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 KIFSANCASCHAGGKNLGVAGKTLKKADLEKYGAYSAMAIGAQVTNGKNAMPAFKGR LKPEEIZBVAAYVLG
 10 20 30 40 50 60 X 70
 KAEAEWK
 80

10. US-08-249-182-3 (1-5)

DBH_BACST DNA-BINDING PROTEIN II (HB) (HU).

ID DBH_BACST STANDARD; PRT; 90 AA.
 AC P02346; P08822;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
 DE DNA-BINDING PROTEIN II (HB) (HU).
 GN HBSU.
 OS BACILLUS STEAROTHERMOPHILUS, BACILLUS CALDOLYTICUS, AND BACILLUS
 OS CALDOTENAX.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.STEAROTHERMOPHILUS, B.CALDOLYTICUS, AND B.CALDOTENAX;
 RM 92354934
 RA PADAS P.M., WILSON K.S., VORGIAS C.E.;
 RL GENE 117:39-44(1992).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=B.STEAROTHERMOPHILUS;
 RM 83160950
 RA KIMURA M., WILSON K.S.;
 RL J. BIOL. CHEM. 258:4007-4011(1983).
 RN [3]
 RP SEQUENCE OF 1-39.
 RC SPECIES=B.CALDOLYTICUS;
 RM 87184910
 RA BECK A., DIJK J., REINHARDT R.;
 RL BIOL. CHEM. HOPPE-SEYLER 368:121-130(1987).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RC SPECIES=B.STEAROTHERMOPHILUS;
 RM 84270702
 RA TANAKA I., APPELT K., DIJK J., WHITE S.W., WILSON K.S.;
 RL NATURE 310:376-381(1984).
 CC -!- FUNCTION: THIS PROTEIN BELONGS OT THE HISTONE LIKE FAMILY OF
 CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING

CC DNA AND TO STABILIZE DNA FROM DENATURATION UNDER EXTREME
 CC ENVIRONMENTAL CONDITIONS.
 CC -!- SUBUNIT: HOMODIMER.
 DR EMBL; M73500; BSHUB1.
 DR EMBL; M73501; BCHUB2.
 DR EMBL; M73502; BCHUB3.
 DR PIR; A02690; DNBS2F.
 DR PIR; A26040; A26040.
 DR PIR; JC1207; JC1207.
 DR PIR; JC1206; JC1206.
 DR PIR; JC1205; JC1205.
 DR PROSITE; PS00045; HISTONE_LIKE.
 KW DNA-BINDING; DNA CONDENSATION.
 FT CONFLICT 9 9 A -> T (IN REF. 3).
 FT CONFLICT 13 13 T -> I (IN REF. 3).
 FT CONFLICT 30 30 D -> E (IN REF. 3).
 SQ SEQUENCE 90 AA; 9716 MW; 38390 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 DAVFDSITEALRKGDQVQLIGFGNFVRRERAARKGRNPQTGEEMEIPASKVPAFKPGKALKDAVK
 30 40 50 60 70 X 80 90

11. US-08-249-182-3 (1-5)

DBH_BACSU DNA-BINDING PROTEIN II (HB) (HU).

ID DBH_BACSU STANDARD; PRT; 92 AA.
 AC P08821;
 DT 01-NOV-1988 (REL. 09, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
 DE DNA-BINDING PROTEIN II (HB) (HU).
 GN HBSU.
 OS BACILLUS SUBTILIS, AND BACILLUS GLOBIGII.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=BACILLUS SUBTILIS;
 RM 92339457
 RA GROCH N., SCHINDELIN H., SCHOLLZ A.S., HAHN U., HEINEMANN U.;
 RL EUR. J. BIOCHEM. 207:677-685(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.SUBTILIS; STRAIN=JH642;
 RM 91216992
 RA MICKA B., GROCH N., HEINEMANN U., MARAHIEL M.A.;
 RL J. BACTERIOL. 173:3191-3198(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.SUBTILIS, AND B.GLOBIGII;
 RM 92354934
 RA PADAS P.M., WILSON K.S., VORGAS C.E.;
 RL GENE 117:39-44(1992).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=B.SUBTILIS;
 RM 87246638
 RA IMBER R., KIMURA M., GROCH N., HEINEMANN U.;
 RL EUR. J. BIOCHEM. 165:547-552(1987).
 CC -!- FUNCTION: THIS PROTEIN BELONGS OT THE HISTONE LIKE FAMILY OF

Gaps = 0 Conservative Substitutions = 0

```

          X  X
          YPAFK
          ||||
MTDLPVVCNRNGAGWWVCGAANGALDSKSRSHLEAETPAFKQSTQHYFFKKQPLPVVESVEEEDDPGVAVENA
      10      20      30      X 40      50      60      70

PSSSEDEENTVEESEKA
      80      90

```

13. US-08-249-182-3 (1-5)
GLN1_METTL NITROGEN FIXATION NIFHD REGION GLNB-LIKE PROTEIN 1

ID GLN1_METTL STANDARD; PRT; 105 AA.
AC P25771;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE NITROGEN FIXATION NIFHD REGION GLNB-LIKE PROTEIN 1 (ORF-105).
OS METHANOCOCCUS THERMOLITHOTROPHICUS.
OC PROKARYOTA; MENDOSICUTES; ARCHAEBACTERIA; METHANOCOCCALES;
OC METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 89343640
RA SOUILLARD N., SIBOLD L.;
RL MOL. MICROBIOL. 3:541-551(1989).
CC -!- FUNCTION: COULD BE INVOLVED IN THE REGULATION OF NITROGEN
CC FIXATION.
CC -!- SIMILARITY: STRONG, TO EUBACTERIAL P(II) PROTEINS.
DR EMBL; X13830; MTNIFHDK.
DR PIR; S06985; S06985.
DR PROSITE; PS00638; PII_GLNB_CTER.
KW TRANSCRIPTION REGULATION; NITROGEN FIXATION.
SQ SEQUENCE 105 AA; 11318 MW; 58675 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

          X  X
          YPAFK
          ||||
MKMIKAIVRPDKVDDIVDSLENAGYPAFKINSVGRGKQGGGLKVGEIFYDELPKTILLIAVNDDEVVVGL
      10      20      X 30      40      50      60      70

IKSSAST

```

14. US-08-249-182-3 (1-5)
YP12_RTBPV HYPOTHETICAL P12 PROTEIN (ORF 2).

ID YP12_RTBPV STANDARD; PRT; 110 AA.
AC P27499;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL P12 PROTEIN (ORF 2).
OS RICE TUNGRO BACILLIFORM VIRUS (ISOLATE PHILIPPINES) (RTBV).
OC VIRIDAE; NOT YET CLASSIFIED.
RN [1]
RP SEQUENCE FROM N.A.
RM 91252204

CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING
 CC DNA AND TO STABILIZE DNA FROM DENATURATION UNDER EXTREME
 CC ENVIRONMENTAL CONDITIONS.
 CC -!- SUBUNIT: HOMODIMER.
 DR EMBL; X66448; BSHBSUHM.
 DR EMBL; X52418; BSHBSV.
 DR EMBL; M73504; BSHUB5.
 DR EMBL; M73503; BGHUB4.
 DR PIR; JC1209; JC1209.
 DR PIR; S00015; S00015.
 DR PIR; S24373; S24373.
 DR PIR; JC1208; JC1208.
 DR PROSITE; PS00045; HISTONE_LIKE.
 KW DNA-BINDING; DNA CONDENSATION.
 FT CONFLICT 40 40 D -> K (IN REF. 4).
 SQ SEQUENCE 92 AA; 9884 MW; 39386 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 YPAFK
 ||||
 DSVFDTILDALKNGDKIQLIGFGNFVRRERSARKGRNPQTGEEIEIPASKVPAFKPGKALKDAVAGK
 30 40 50 60 70 X 80 90

12. US-08-249-182-3 (1-5)
 PINO_ECOLI PINO PROTEIN.

ID PINO_ECOLI STANDARD; PRT; 105 AA.
 AC P03825;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE PINO PROTEIN.
 GN PINO.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 82137054
 RA OLINS P.O., NOMURA M.;
 RL CELL 26:205-211(1981).
 RN [2]
 RP CHARACTERIZATION.
 RM 92021770
 RA GUZMAN E.C., JIMENEZ-SANCHEZ A.;
 RL RES. MICROBIOL. 142:137-140(1991).
 RN [3]
 RP IDENTIFICATION OF PROTEIN.
 RM 92041646
 RA GUZMAN E.C., JIMENEZ-SANCHEZ A.;
 RL J. BACTERIOL. 173:7409-7409(1991).
 CC -!- FUNCTION: CALCIUM-BINDING PROTEIN THAT MAY BE REQUIRED FOR THE
 CC INITIATION OF CHROMOSOME REPLICATION.
 DR EMBL; V00344; ECRPSL.
 DR PIR; A04450; Q0ECRP.
 DR ECGENE; EG11263; PINO.
 KW CALCIUM-BINDING; DNA REPLICATION.
 SQ SEQUENCE 105 AA; 11692 MW; 52846 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.39
 Residue Identity = 80% Matches = 4 Mismatches = 1

RC STRAIN=Y1140;
 RM 89034156
 RA CHANG Y.-D., DICKSON R.C.;
 RL J. BIOL. CHEM. 263:16696-16703(1988).
 CC -!- FUNCTION: LAC12 MEDIATES THE TRANSPORT OF LACTOSE AND IT WOULD
 CC APPEAR THAT THE PERMEASE WORKS IN PART BY A PROTON SYMPORT
 CC MECHANISM.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; X06997; KLLAC12.
 DR PIR; A31776; A31776.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2.
 KW DUPLICATION; TRANSMEMBRANE; TRANSPORT; SUGAR TRANSPORT; SYMPORT.
 SQ SEQUENCE 587 AA; 65383 MW; 1854896 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                ||| | ||
SSSSSLQKKPINTIEHKDTLGNDRDHKEALNSDNDNTSGLKINGVPIEDAREEVLLPGYLSKQYYKLYGLCF
  10      20      30      40      50      X  60      X  70

ITYLCATMGGYDGMGSIYTEDAYLKYYHLDINSSSG
  80      90     100     110
  
```

9. US-08-249-182-5 (1-10)

RRP1_DROME RECOMBINATION REPAIR PROTEIN 1 (DNA-(APURINIC OR A

ID RRP1_DROME STANDARD; PRT; 679 AA.
 AC P27864;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE RECOMBINATION REPAIR PROTEIN 1 (DNA-(APURINIC OR APYRIMIDINIC SITE)
 DE LYASE) (EC 4.2.99.18).
 GN RRP1.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RM 91319767
 RA SANDER M., LOWENHAUPT K., RICH A.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:6780-6784(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RM 91360356
 RA SANDER M., LOWENHAUPT K., LANE W.S., RICH A.;
 RL NUCLEIC ACIDS RES. 19:4523-4529(1991).
 CC -!- FUNCTION: COULD PROMOTE HOMOLOGOUS RECOMBINATION AT SITES OF DNA
 CC DAMAGE. RRP1 HAS APURINIC ENDONUCLEASE AND DOUBLE-STRANDED DNA 3'
 CC EXONUCLEASE, ACTIVITIES AND CARRIES OUT SINGLE-STRANDED DNA
 CC RENATURATION IN A MG(2+)-DEPENDENT MANNER.
 CC -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE NEAR APURINIC OR
 CC APYRIMIDINIC SITES TO PRODUCTS WITH 5'-PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: BELONGS TO THE AP/EXO A FAMILY OF DNA REPAIR ENZYMES.
 DR EMBL; M62472; DMRRP1.
 DR FLYBASE; 04584; RRP1.
 DR PIR; S28366; S28366.

PI CONFLICT 278 278 A -7 P (IN REF. 27).
SQ SEQUENCE 323 AA; 34973 MW; 594373 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
LDPVFTPATGTPVVGGLSYREGLYITEEIKTGLLSGLDIMEVNPTLGKTPEEVTRTVNTAVALTLSCFGTK
 240      250      260      270      280      X 290      X 300

REGNHKPPETDYLKPPK
 310      320
```

7. US-08-249-182-5 (1-10)
VC13_SFVKA PROTEIN C13.

ID VC13_SFVKA STANDARD; PRT; 500 AA.
AC P32206;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE PROTEIN C13.
GN C13L.
OS SHOPE FIBROMA VIRUS (STRAIN KASZA) (SFV).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC LEPORIPOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RA MASSUNG R.F., JAYARAMA V., MOYER R.W.;
RL SUBMITTED (XXX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: TO VACCINIA VIRUS ORF F3.
DR EMBL; L22013; PXSHPHLSB.
SQ SEQUENCE 500 AA; 57475 MW; 1384311 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                ||| |||
TDILLSLSLFDLRIILKSGELDVSSDDVLLFIIKWSRHKKSNRRKSFTLVTEVLRNYLSIYGKYKLTWKL
 160      170      180      190      200      X 210      X 220

ARFGKNNNVELNENELPRISYQHRFTNRRYTMVTPSSF
 230      240      250      260
```

8. US-08-249-182-5 (1-10)
LACP_KLULA LACTOSE PERMEASE.

ID LACP_KLULA STANDARD; PRT; 587 AA.
AC P07921;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE LACTOSE PERMEASE.
GN LAC12.
OS KLUYVEROMYCES LACTIS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.

DR PIR; S02132; S02132.
 DR PIR; A26370; A26370.
 DR MIM; 207800; TENTH EDITION.
 DR PROSITE; PS00147; ARGINASE_1.
 DR PROSITE; PS00148; ARGINASE_2.
 KW UREA CYCLE; ARGININE METABOLISM; HYDROLASE; MAGNESIUM.
 SQ SEQUENCE 322 AA; 34734 MW; 574214 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
LDPSTPATGTPVVGGLTYREGLYITEEIIYKTGLLSGLDIMEVNP SLGKTPEEVTRTVNTAVAITLACFGLA
 240      250      260      270      280      X 290      X 300

REGNHKPIDYLNPPK
310      320

```

6. US-08-249-182-5 (1-10)

ARGI_RAT ARGINASE (EC 3.5.3.1).

ID ARGI_RAT STANDARD; PRT; 323 AA.
 AC P07824;
 DT 01-AUG-1988 (REL. 08, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE ARGINASE (EC 3.5.3.1).
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 DC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RM 88115364
 RA OHTAKE A., TAKIGUCHI M., SHIGETO Y., AMAYA Y., KAWAMOTO S., MORI M.;
 RL J. BIOL. CHEM. 263:2245-2249(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RM 87194847
 RA KAWAMOTO S., AMAYA Y., MURAKAMI K., TOKUNAGA F., IWANAGA S.,
 RA KOBAYASHI K., SAHEKI T., KIMURA S., MORI M.;
 RL J. BIOL. CHEM. 262:6280-6283(1987).
 CC -!- CATALYTIC ACTIVITY: L-ARGININE + H(2)O = L-ORNITHINE + UREA.
 CC -!- PATHWAY: FIRST STEP IN ARGININE DEGRADATION IN THE UREA CYCLE.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- COFACTOR: MN(2+).
 CC -!- INDUCTION: BY ARGININE OR HOMODARGININE.
 DR EMBL; M17924; RNARG1.
 DR EMBL; M17925; RNARG2.
 DR EMBL; M17926; RNARG3.
 DR EMBL; M17927; RNARG4.
 DR EMBL; M17928; RNARG5.
 DR EMBL; M17929; RNARG6.
 DR EMBL; M17930; RNARG7.
 DR EMBL; M17931; RNARG8.
 DR EMBL; J02720; RNARGL.
 DR PIR; A26702; A26702.
 DR PIR; A28358; A28358.
 DR PROSITE; PS00147; ARGINASE_1.
 DR PROSITE; PS00148; ARGINASE_2.
 KW UREA CYCLE; ARGININE METABOLISM; HYDROLASE; MAGNESIUM.

DC PROKARYOTA; GRACILICUTES; SCOTUBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 93268094
 RA COLLINS C.M., GUTMAN D.M., LAMAN H.;
 RL MOL. MICROBIOL. 8:187-198(1993).
 CC -!- FUNCTION: PROBABLY FACILITATING NICKEL INCORPORATION.
 DR EMBL; L07039; KPURE.
 DR PIR; S32937; S32937.
 KW NICKEL.
 SQ SEQUENCE 270 AA; 29953 MW; 374889 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 ||| | ||
 MLPPLKKGWQRTLDLRFQAGGKTVLASAQHVGLTVQRPFYEEETCHLYLLHPPGGIVGGDELTISAHLA
 10 20 30 40 X 50 X 60 70
 PGCHTLITMPGASKFYRSSGAQALVRQQLT
 80 90 100

5. US-08-249-182-5 (1-10)
 ARG1_HUMAN ARGINASE (EC 3.5.3.1).

ID ARG1_HUMAN STANDARD; PRT; 322 AA.
 AC P05089;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE ARGINASE (EC 3.5.3.1).
 GN ARG1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RM 87092419
 RA HARAGUCHI Y., TAKIGUCHI M., AMAYA Y., KAWAMOTO S., MATSUDA I.,
 RA MORI M.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 84:412-415(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RM 89016562
 RA TAKIGUCHI M., HARAGUCHI Y., MORI M.;
 RL NUCLEIC ACIDS RES. 16:8789-8802(1988).
 CC -!- CATALYTIC ACTIVITY: L-ARGININE + H(2)O = L-ORNITHINE + UREA.
 CC -!- PATHWAY: FIRST STEP IN ARGININE DEGRADATION IN THE UREA CYCLE.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- COFACTOR: MN(2+).
 CC -!- INDUCTION: BY ARGININE OR HOMOARGININE.
 CC -!- DISEASE: DEFICIENCY IN ARG1 IS THE CAUSE OF ARGININEMIA.
 DR EMBL; M14502; HSARGL.
 DR EMBL; X12662; HSARG1.
 DR EMBL; X12663; HSARG2.
 DR EMBL; X12664; HSARG3.
 DR EMBL; X12665; HSARG4.
 DR EMBL; X12667; HSARG6.
 DR EMBL; X12668; HSARG7.

Initial Score = 6 Optimized Score = 6 Significance = 4.65
Residue Identity = 75% Matches = 6 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||| ||
AVTVDR LGNMA SASTPVT LAMFWPDI QPGQRVL VLT YGSGATWGAAL YRKPEEVN RPC
    180      190      200      210      220 X      X

```

3. US-08-249-182-5 (1-10)

UNG_HCMVA URACIL-DNA GLYCOSYLASE (EC 3.2.2.-).

ID UNG_HCMVA STANDARD; PRT; 250 AA.
AC P16769;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE URACIL-DNA GLYCOSYLASE (EC 3.2.2.-).
GN UL114.
OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 90269039
RA CHEE M.S., BANKIER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R.,
RA HORSNELL T., HUTCHISON C.A. III, KOUZARIDES T., MARTIGNETTI J.A.,
RA PREDDIE E., SATCHWELL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;
RL CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990).
CC -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE
CC AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA
CC POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE.
CC -!- SIMILARITY: STRONGLY CONSERVED IN ALL SPECIES.
DR EMBL; X17403; HEHCMVCG.
DR PIR; S09881; DGBEL5.
DR PROSITE; PS00130; U_DNA_GLYCOSYLASE.
KW DNA REPAIR; HYDROLASE; GLYCOSIDASE.
SQ SEQUENCE 250 AA; 28353 MW; 307699 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||| ||
LLTPDEGARVFCL SADWIRFLSLPDHDTVLLRDTVA AVEGARQLEMVYP APEHVHRWSYLC PPEQVRVVI VG
    20      30      40      50      60      70      X 80

GDPYCDGSASGLAFGTL AGRPPPPSLNNVFRELARTVD
    90      100      110      120

```

4. US-08-249-182-5 (1-10)

URED_KLEPN UREASE OPERON URED PROTEIN.

ID URED_KLEPN STANDARD; PRT; 270 AA.
AC 002944;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE UREASE OPERON URED PROTEIN.
GN URED.
OS KLEBSIELLA PNEUMONIAE.

38. PCED_ECOLI	HYPOTHETICAL 17.3 KD PROTEIN	173	5	5	3.49	0
39. AROL_ECOLI	SHIKIMATE KINASE II (EC 2.7.1	174	5	5	3.49	0
40. GLUC_OCTDE	GLUCAGON PRECURSOR.	180	5	6	3.49	0

1. US-08-249-182-5 (1-10)

EXON_HSVSA ALKALINE EXONUCLEASE (3.1.11.-).

ID EXON_HSVSA STANDARD; PRT; 483 AA.
AC 001013;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE ALKALINE EXONUCLEASE (3.1.11.-).
GN 37.
OS HERPESVIRUS SAIMIRI (STRAIN 11).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 92333688
RA ALBRECHT J.-C., NICHOLAS J., BILLER D., CAMERON K.R., BIESINGER B.,
RA NEWMAN C., WITTMANN S., CRAXTON M.A., COLEMAN H., FLECKENSTEIN B.,
RA HONESS R.W.;
RL J. VIROL. 66:5047-5058(1992).
DR EMBL; X64346; HSGEND.
DR PIR; H36809; Q0BEN4.
KW HYDROLASE; NUCLEASE; EXONUCLEASE.
SQ SEQUENCE 483 AA; 55554 MW; 1295833 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.82
Residue Identity = 70% Matches = 7 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

					X	X
					PEEVTRPNYL	
GGFIGNEKIGTYFSKNVCNNIAAKGVPKLADVYKACEKMNLRQ0SEICLLIEVTRGQYLNSLWDALRDGTI						
70	80	90	100	110	120	130
SSSKFYWATKKQNSTKKIFEPWPIKNDYVYVAGPLAFGL						
140	150	160	170			

2. US-08-249-182-5 (1-10)

YPH1_PSEAE HYPOTHETICAL PROTEIN IN PHNA 5'REGION (ORF1) (FRAG

ID YPH1_PSEAE STANDARD; PRT; 229 AA.
AC P20582;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN IN PHNA 5'REGION (ORF1) (FRAGMENT).
OS PSEUDOMONAS AERUGINOSA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC PSEUDOMONADACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RM 90130326
RA ESSAR D.W., EBERLY L., HADERO A., CRAWFORD I.P.;
RL J. BACTERIOL. 172:884-900(1990).
DR EMBL; M33810; PAPHNABA.
DR PIR; D35116; D35116.
KW HYPOTHETICAL PROTEIN.
FT NON_TER 1 1
SQ SEQUENCE 229 AA; 24636 MW; 257164 CN;

0.86

00:00:55.00

3826 .

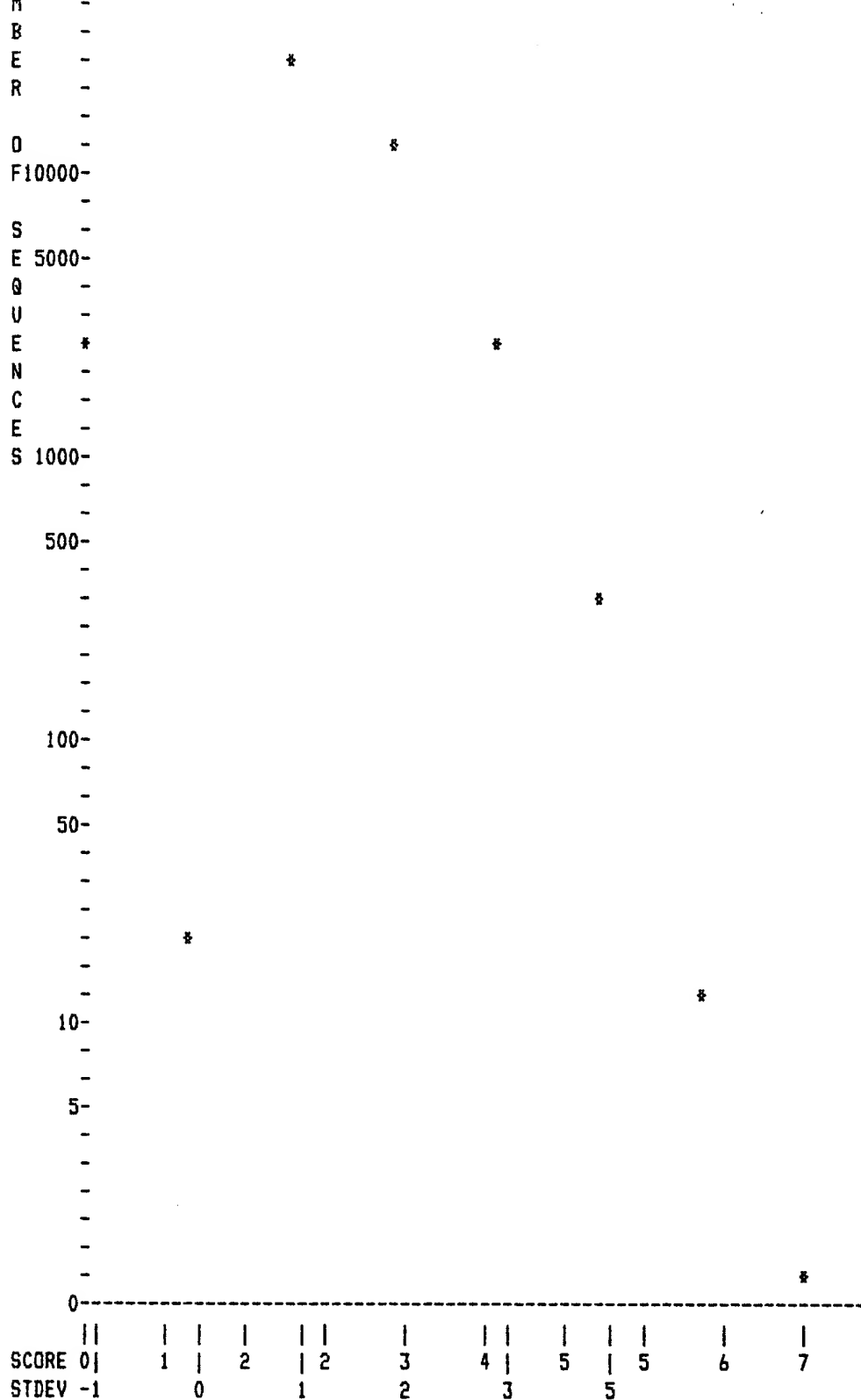
Cut-off raised to 5.

Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

37. VS11 BOTGA	NONSTRUCTURAL PROTEIN.	170	5	5	3.49	0
----------------	------------------------	-----	---	---	------	---



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

DATE 28-May-1986 #sequence_revision 28-May-1986 #text_change
 03-Feb-1994
 ACCESSIONS A00604
 REFERENCE A00604
 #authors Poorman, R.A.; Randolph, A.; Kemp, R.G.; Heinrikson, R.L.
 #journal Nature (1984) 309:467-469
 #title Evolution of phosphofructokinase-gene duplication and
 creation of new effector sites.
 #cross-references MUID:84219739
 #accession A00604
 ##molecule_type protein
 ##residues 1-749 ##label P00
 COMMENT The active enzyme, a tetramer of identical chains, catalyzes the
 key control step of glycolysis, the phosphorylation of
 fructose-6-phosphate by ATP to form fructose-1,6-bisphosphate in
 the presence of magnesium ion. It is an allosteric enzyme
 activated by ADP, AMP, or fructose bisphosphate and inhibited by
 ATP or citrate.
 CLASSIFICATION #superfamily human 6-phosphofructokinase;
 6-phosphofructokinase 1 homology
 KEYWORDS acetylation; allosteric regulation; duplication; glycolysis;
 phosphoprotein; phosphotransferase
 FEATURE
 17-326 #domain 6-phosphofructokinase 1 homology #label 6PF1\
 402-658 #domain 6-phosphofructokinase 1 homology #label 6PF2\
 1 #modified_site acetylated amino end (Thr) #status
 experimental\
 744 #binding_site phosphate (Ser) (covalent) #status
 experimental
 SUMMARY #length 749 #molecular-weight 81839 #checksum 3199
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 || || ||
 GRLRAAHNLVKRGITNLCVIGDGLTGADTFRSEWSDLLSLQKAGKITAEATRSSLNIVGLVGSIDND
 100 110 120 130 140 X 150 X 160

FCGTDMTIGTDSALHRITEIVDAITTTA@SHQRTFVLE
 170 180 190 200

> 0 <
 0| 10 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_5s.res made by on Thu 22 Sep 94 10:21:04-PDT.

Query sequence being compared:US-08-249-182-5 (1-10)
 Number of sequences searched: 36000
 Number of scores above cutoff: 3826

Results of the initial comparison of US-08-249-182-5 (1-10) with:
 Data bank : Swiss-Prot 28, all entries

100000-
 -
 N -
 U50000-

SUMMARY #length 677 #molecular-weight 74882 #checksum 176
SEQUENCE

Initial Score = 6 Optimized Score = 7 Significance = 4.11
Residue Identity = 72% Matches = 8 Mismatches = 2
Gaps = 1 Conservative Substitutions = 0

```

                                X      X
                                PEEVTR-PNYL
                                ||||| ||
PADKEFNLKICSNWNVAGLRAWLKKDGLQLIDLEEPDIFCLGETKCANDQLPEEVTRLPGYHPYWLCPGGYA
420      430      440      450      460      470      480      490

GVAIYSKIMPIHVEYGIGNEEFDDVGRMITAEYEKFYLI
      500      510      520      530
```

14. US-08-249-182-5 (1-10)

A48832 gp138=cell surface glycoprotein - slime mold (Dict

ENTRY A48832 #type complete
TITLE gp138=cell surface glycoprotein - slime mold (Dictyostelium
discoideum)
ORGANISM #formal_name Dictyostelium discoideum
DATE 01-Dec-1993; #sequence_revision 01-Dec-1993; #text_change
01-Dec-1993
ACCESSIONS A48832
REFERENCE A48832
#authors Fang, H.; Higa, M.; Suzuki, K.; Aiba, K.; Urushihara, H.;
Yanagisawa, K.
#journal Dev. Biol. (1993) 156:201-208
#title Molecular cloning and characterization of two genes encoding
gp138, a cell surface glycoprotein involved in the sexual
cell fusion of Dictyostelium discoideum.
#cross-references MUID:93193972
#contents AX-3 cells
#accession A48832
##status preliminary
##molecule_type DNA; protein
##residues 1-730 ##label FAN
##cross-references NCBIN:127258; NCBIP:127259
##note sequence extracted from NCBI backbone
SUMMARY #length 730 #molecular-weight 80960 #checksum 7716
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                || | |||
NTFSYECIDFSNIADFSDDYNDYEGYLPNIDNAPLLTGVDIYQSVVVGDIPESYCRINYLSLYYNQLNGTVP
      330      340      350      360      370      X 380      X 390

SCIQCLGGVKGGDIVLPNPFLNFKTTEPYCPTFKIDE
      400      410      420      430
```

15. US-08-249-182-5 (1-10)

KIRBF 6-phosphofructokinase (EC 2.7.1.11) - rabbit

ENTRY KIRBF #type complete
TITLE 6-phosphofructokinase (EC 2.7.1.11) - rabbit
ALTERNATE_NAMES phosphofructokinase 1; phosphohexokinase
ORGANISM #formal_name Oryctolagus cuniculus #common name domestic

12. US-08-249-182-5 (1-10)

A31776 lactose permease - yeast (*Kluyveromyces marxianus*)

ENTRY A31776 #type complete
TITLE lactose permease - yeast (*Kluyveromyces marxianus* var. *lactis*)
ORGANISM #formal_name *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*
DATE 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
ACCESSIONS A31776
REFERENCE A92683
#authors Chang, Y.D.; Dickson, R.C.
#journal J. Biol. Chem. (1988) 263:16696-16703
#title Primary structure of the lactose permease gene from the yeast *Kluyveromyces lactis*. Presence of an unusual transcript structure.
#cross-references MUID:89034156
#accession A31776
##molecule_type DNA
##residues 1-587 ##label CHA
##cross-references GB:X06997
SUMMARY #length 587 #molecular-weight 65383 #checksum 2470
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
||| | ||
SSSSSLQKKPINTIEHKDTLGNDRDHKEALNSDNDNTSGLKINGVPIEDAREEVLLPGYLSKQYYKLYGLCF
10 20 30 40 50 X 60 X 70

ITYLCATMGGYDGMGSIYTEDAYLKYYHLDINSSSG
80 90 100 110

13. US-08-249-182-5 (1-10)

S28366 recombination repair protein 1 - fruit fly (*Drosophila*)

ENTRY S28366 #type complete
TITLE recombination repair protein 1 - fruit fly (*Drosophila melanogaster*)
ORGANISM #formal_name *Drosophila melanogaster*
DATE 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Jun-1993
ACCESSIONS S28366
REFERENCE S28366
#authors Sander, M.; Lowenhaupt, K.; Lane, W.S.; Rich, A.
#journal Nucleic Acids Res. (1991) 19:4523-4529
#title Cloning and characterization of *Rrp1*, the gene encoding *Drosophila* strand transferase: carboxy-terminal homology to DNA repair endo/exonucleases.
#accession S28366
##molecule_type mRNA
##residues 1-679 ##label SAN
##cross-references EMBL:M62472
GENETICS
#gene *Rrp1*

TITLE arginase (EC 3.5.3.1), hepatic - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change
 28-Apr-1993
 ACCESSIONS A28358
 REFERENCE A28358
 #authors Ohtake, A.; Takiguchi, M.; Shigeto, Y.; Anaya, Y.; Kawamoto,
 S.; Mori, M.
 #journal J. Biol. Chem. (1988) 263:2245-2249
 #title Structural organization of the gene for rat liver-type
 arginase.
 #cross-references MUID:88115364
 #accession A28358
 ##molecule_type DNA
 ##residues 1-323 ##label OHT
 ##cross-references GB:M17924
 KEYWORDS hydrolase; urea cycle
 SUMMARY #length 323 #molecular-weight 34973 #checksum 6211
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
LDPVFTPATGTPVVGGLSYREGLYITEEIKTGLLSGLDIMEVNPTLGKTPEEVTRTVNTAVALTLSCFGTK
  240      250      260      270      280      X 290      X 300

REGNHKPETDYLKPPK
  310      320
  
```

11. US-08-249-182-5 (1-10)
 A26702 arginase (EC 3.5.3.1), hepatic - rat

ENTRY A26702 #type complete
 TITLE arginase (EC 3.5.3.1), hepatic - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change
 28-Apr-1993
 ACCESSIONS A26702
 REFERENCE A26702
 #authors Kawamoto, S.; Anaya, Y.; Murakami, K.; Tokunaga, F.; Iwanaga,
 S.; Kobayashi, K.; Saheki, T.; Kimura, S.; Mori, M.
 #journal J. Biol. Chem. (1987) 262:6280-6283
 #title Complete nucleotide sequence of cDNA and deduced amino acid
 sequence of rat liver arginase.
 #cross-references MUID:87194847
 #accession A26702
 ##molecule_type mRNA
 ##residues 1-323 ##label KAW
 KEYWORDS hydrolase; urea cycle
 SUMMARY #length 323 #molecular-weight 34999 #checksum 6406
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
LDPVFTPATGTPVVGGLSYREGLYITEEIKTGLLSGLDIMEVNPTLGKTPEEVTRTVNTAVPLTLSCFGTK
  240      250      260      270      280      X 290      X 300
  
```

#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:412-415
 #title Molecular cloning and nucleotide sequence of cDNA for human
 liver arginase.
 #cross-references MUID:87092419
 #accession A26370
 ##molecule_type mRNA
 ##residues 1-322 ##label HAR
 KEYWORDS hydrolase; urea cycle
 SUMMARY #length 322 #molecular-weight 34734 #checksum 2748
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 |||||
 LDPSFTPATGTPVVGGLTYREGLYITEEYKTLGLSGLDIMEVNPGLKTPEEVTRTVNTAVAILLACFGLA
 240 250 260 270 280 X 290 X 300
 REGNHKPIDYLNPPK
 310 320

9. US-08-249-182-5 (1-10)

B26370 arginase (EC 3.5.3.1), hepatic - rat

ENTRY B26370 #type complete
 TITLE arginase (EC 3.5.3.1), hepatic - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change
 28-Apr-1993
 ACCESSIONS B26370
 REFERENCE A94160
 #authors Haraguchi, Y.; Takiguchi, M.; Anaya, Y.; Kawamoto, S.;
 Matsuda, I.; Mori, M.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:412-415
 #title Molecular cloning and nucleotide sequence of cDNA for human
 liver arginase.
 #cross-references MUID:87092419
 #accession B26370
 ##molecule_type protein
 ##residues 1-323 ##label HAR
 KEYWORDS hydrolase; urea cycle
 SUMMARY #length 323 #molecular-weight 34927 #checksum 6462
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 |||||
 LDPVFTPATGTPVVGGLSYRGGLYITEEYKTLGLSGLDIMEVNPTLGKTPEEVTRTVNTAVPLTLSCFGTK
 240 250 260 270 280 X 290 X 300
 REGNHKPETDYLKPPK
 310 320

10. US-08-249-182-5 (1-10)

A28358 arginase (EC 3.5.3.1), hepatic - rat

ENTRY A28358 #type complete

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                ||| | ||
MLPPLKKGWQATLDRFHQAGGKTVLASAQHVGPLTVQRPFYEEETCHLYLLHPPGGIVGGDELTISAHLA
      10      20      30      40 X      50 X      60      70

PGCHTLITMPGASKFYRSSGAQALVRQQLT
      80      90     100
```

7. US-08-249-182-5 (1-10)

S02132 arginase (EC 3.5.3.1), hepatic - human

ENTRY S02132 #type complete
TITLE arginase (EC 3.5.3.1), hepatic - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
28-Apr-1993
ACCESSIONS S02132
REFERENCE S02132
#authors Takiguchi, M.; Haraguchi, Y.; Mori, M.
#journal Nucleic Acids Res. (1988) 16:8789-8802
#title Human liver-type arginase gene: structure of the gene and
analysis of the promoter region.
#cross-references MUID:89016562
#accession S02132
##molecule_type DNA
##residues 1-322 ##label TAK
##cross-references EMBL:X12662
GENETICS
#introns 19/3; 44/1; 102/2; 155/3; 187/2; 222/2; 268/1
KEYWORDS hydrolase; urea cycle
SUMMARY #length 322 #molecular-weight 34735 #checksum 2400
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                PEEVTRPNYL
                                |||||
LDPSFTPATGTPVVGGGLTYREGLYITEEIKTGLLSGLDINEVNPSLGKTPEEVTRTVNTAVAITLACFGLA
      240      250      260      270      280      X 290      X 300

REGNHKPIDYLNPPK
      310      320
```

8. US-08-249-182-5 (1-10)

A26370 arginase (EC 3.5.3.1), hepatic - human

ENTRY A26370 #type complete
TITLE arginase (EC 3.5.3.1), hepatic - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change
28-Apr-1993
ACCESSIONS A26370
REFERENCE A94160
#authors Haraguchi, Y.; Takiguchi, M.; Anaya, Y.; Kawamoto, S.;
Matsuda, I.; Mori, M.

90 100 110 120

5. US-08-249-182-5 (1-10)

S32937 ureD protein - Klebsiella pneumoniae

```
ENTRY          S32937          #type complete
TITLE          ureD protein - Klebsiella pneumoniae
ORGANISM       #formal_name Klebsiella pneumoniae
DATE           31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
              31-Dec-1993
ACCESSIONS     S32937
REFERENCE      S32937
#authors       Collins, C.M.; Gutman, D.M.; Lanan, H.
#journal       Mol. Microbiol. (1993) 8:187-198
#title         Identification of a nitrogen-regulated promoter controlling
              expression of Klebsiella pneumoniae urease genes.
#accession     S32937
##molecule_type DNA
##residues     1-270 ##label COL
##cross-references EMBL:L07039
```

GENETICS

```
#gene      ureD
```

```
#start codon  GTG
```

SUMMARY #length 270 #molecular-weight 29953 #checksum 5656

SEQUENCE

```
Initial Score      =      6  Optimized Score =      6  Significance =  4.11
Residue Identity  =     60%  Matches          =      6  Mismatches   =      4
Gaps              =      0  Conservative Substitutions =      0
```

X 10
PEEVTRPNYL
||| | ||
MLPPLKKGWQRTLDLRFQAGGKTVLASAQHVGPLTVQRPFYEEETCHLYLLHPPGGIVGGDELTI SAHLA
10 20 30 40 X 50 X 60 70
PGCHTLITMPGASKFYRSSGAQALVRQQLT
80 90 100

6. US-08-249-182-5 (1-10)

A42887 urease - *Klebsiella pneumoniae*

```
ENTRY          A42887      #type complete
TITLE          urease - Klebsiella pneumoniae
ORGANISM       #formal_name Klebsiella pneumoniae
DATE           31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
                31-Dec-1993
ACCESSIONS     A42887
REFERENCE      A42887
#authors       Lee, M.H.; Mulrooney, S.B.; Renner, M.J.; Markowicz, Y.;
                Hausinger, R.P.
#journal       J. Bacteriol. (1992) 174:4324-4330
#title         Klebsiella aerogenes urease gene cluster: sequence of ureD
                and demonstration that four accessory genes (ureD, ureE,
                ureF, and ureG) are involved in nickel metallocenter
                biosynthesis.
#accession     A42887
##status       preliminary
##molecule_type DNA
##residues     1-270 ##label LEE
##cross-references GB:M55391
```

SUMMARY

SEQUENCE

evolutionary implications.

#cross-references MUID:90130326
#accession D35116
##status preliminary
##molecule_type DNA
##residues 1-229 ##label ESS
##cross-references GB:M33810
SUMMARY #length 229 #checksum 9355
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 75% Matches = 6 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
||||||
AVTVDR LGNMASASTPVT LAMFWDI QPGQRVL VLT YGSGATWGAAL YRKPEEVNRPC
180 190 200 210 220 X X

4. US-08-249-182-5 (1-10)

DGBEL5 uracil-DNA glycosylase (EC 3.2.2.-) - human cytome

ENTRY DGBEL5 #type complete
TITLE uracil-DNA glycosylase (EC 3.2.2.-) - human cytomegalovirus
(strain AD169)
ALTERNATE_NAMES UL114 protein
ORGANISM #formal_name human cytomegalovirus, human herpesvirus 5
#note host Homo sapiens (man)
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
24-Feb-1994
ACCESSIONS S09881
REFERENCE S09749
#authors Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.;
Cerny, R.; Horsnell, T.; Hutchison III, C.A.; Kouzarides,
T.; Martignetti, J.A.; Preddie, E.; Satchwell, S.C.;
Tonlinson, P.; Weston, K.M.; Barrell, B.G.
#journal Curr. Top. Microbiol. Immunol. (1990) 154:125-169
#title Analysis of the protein-coding content of the sequence of
human cytomegalovirus strain AD169.
#cross-references MUID:90269039
#accession S09881
##molecule_type DNA
##residues 1-250 ##label CHE
##cross-references EMBL:X17403
##note possible protein-coding frames are given in this paper
##note neither protein nor nucleic acid sequence is given in
this paper
##note the DNA sequence was submitted to EMBL, December 1989,
in computer-readable form
CLASSIFICATION #superfamily uracil-DNA glycosylase
KEYWORDS glycosidase; hydrolase
SUMMARY #length 250 #molecular-weight 28353 #checksum 6234
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.11
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
||||||
LLTPDEGARVFCLSADWIRFLSPDHDTVLLRDTVAAVEGARQLEHVYPAPEHVHRWSYLCPPEQVRVVIVG
20 30 40 50 60 70 X 80

2. US-08-249-182-5 (1-10)

QQBEN4 alkaline exonuclease (EC 3.1.11.-) - saimiriine he

ENTRY QQBEN4 #type complete
 TITLE alkaline exonuclease (EC 3.1.11.-) - saimiriine herpesvirus 1
 (strain 11)
 ORGANISM #formal_name saimiriine herpesvirus 1
 #note host Saimiri sciureus (common squirrel monkey)
 DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change
 04-Mar-1994
 ACCESSIONS H36809
 REFERENCE A36806
 #authors Albrecht, J.
 #submission submitted to the EMBL Data Library, January 1992
 #description Primary structure of the herpesvirus saimiri genome.
 #accession H36809
 ##molecule_type DNA
 ##residues 1-483 ##label ALB
 ##cross-references GB:X64346
 REFERENCE A37309
 #authors Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.;
 Biesinger, B.; Newman, C.; Wittmann, S.; Craxton, M.A.;
 Coleman, H.; Fleckenstein, B.; Honess, R.W.
 #journal J. Virol. (1992) 66:5047-5058
 #title Primary structure of the herpesvirus saimiri genome.
 #cross-references MUID:92333688
 #contents annotation; possible protein-coding frames
 #note neither protein nor nucleotide sequence is given in this
 paper
 GENETICS
 #gene 37
 CLASSIFICATION #superfamily human cytomegalovirus alkaline exonuclease
 KEYWORDS exonuclease; hydrolase
 SUMMARY #length 483 #molecular-weight 55554 #checksum 6775
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 5.13
 Residue Identity = 70% Matches = 7 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

X X
 PEEVTRPNYL

||||| ||

GGFIGNEKIGTYFSKNVCNNIAAKGVPLADVYKACEKMNLRQSEICLLIEVTRGQYLNSLWDALRDGTI

70 80 90 100 110 120 130

SSSKFYWATKKQNSTKKIFEPWPIKNDYYVAGPLAFGL

140 150 160 170

3. US-08-249-182-5 (1-10)

D35116 hypothetical protein 2 (phnA 5' region) - Pseudomo

ENTRY D35116 #type fragment
 TITLE hypothetical protein 2 (phnA 5' region) - Pseudomonas
 aeruginosa (fragment)
 ORGANISM #formal_name Pseudomonas aeruginosa
 DATE 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change
 30-Sep-1993
 ACCESSIONS D35116
 REFERENCE A35116
 #authors Essar, D.W.; Eberly, L.; Hadero, A.; Crawford, I.P.
 #journal J. Bacteriol. (1990) 172:884-900
 #title Identification and characterization of genes for a second
 anthranilate synthase in Pseudomonas aeruginosa:
 interchangeability of the two anthranilate synthases and

17. S30873	paraspornal crystal protein -	1228	8	8	4.11	0
**** 3 standard deviations above mean ****						
20. S13898	alkaline phosphatase (EC 3.1.	16	5	5	3.08	0
21. S34480	hypothetical protein - Synech	19	5	5	3.08	0
22. C35646	mast cell proteinase MMCP-5 (30	5	6	3.08	0
23. S19295	prolactin - rat	42	5	5	3.08	0
24. A41256	protein-tyrosine kinase (EC 2	50	5	5	3.08	0
25. S40660	rpoC1 protein - maize	54	5	5	3.08	0
26. T2NJ2Y	short toxin CM-2 - Egyptian c	61	5	5	3.08	0
27. A43732	BR6 protein - midge (Chironom	79	5	5	3.08	0
28. PH1048	Ig light chain V region (clon	81	5	5	3.08	0
29. C34435	hypothetical protein IR-II (r	90	5	5	3.08	0
30. S33718	Ribosomal protein S7 - Sulfol	94	5	6	3.08	0
31. S26259	T-cell receptor Vbeta 6.7 - H	96	5	5	3.08	0
32. W6MLB4	E6 protein - bovine papilloma	99	5	5	3.08	0
33. PH1046	Ig light chain V region (clon	101	5	5	3.08	0
34. S03492	T-cell receptor beta chain V-	102	5	5	3.08	0
35. PH1052	Ig light chain V region (clon	103	5	5	3.08	0
36. PH1051	Ig light chain V region (clon	103	5	5	3.08	0
37. PH1050	Ig light chain V region (clon	103	5	5	3.08	0
38. PH1049	Ig light chain V region (clon	103	5	5	3.08	0
39. PH1047	Ig light chain V region (clon	103	5	5	3.08	0
40. PH1104	Ig light chain V region (clon	104	5	5	3.08	0

1. US-08-249-182-5 (1-10)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
TITLE autotaxin - human (fragments)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change 08-May-1993
ACCESSIONS A42329
REFERENCE A42329
#authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann, E.; Liotta, L.A.
#journal J. Biol. Chem. (1992) 267:2524-2529
#title Identification, purification, and partial sequence analysis of autotaxin, a novel motility-stimulating protein.
#cross-references MUID:92129337
#accession A42329
##status preliminary
##molecule_type protein
##residues 1-114 ##label STR
##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518; NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510; NCBIP:78509; NCBIP:78508; NCBIP:78503
##note sequence extracted from NCBI backbone
SUMMARY #length 114 #checksum 7335
SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 7.19
Residue Identity = 90% Matches = 9 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X X
PEEVTRPNYL
||||| ||||
DFFRVNSMGTVFVGYGPTFKGGQPLWITATKSPFFENINLYYDVPWNETIPEEVTPNYLQAEVSYPAFKPX
40 50 60 70 80 X 90 100

LDVYKWHVAAN
110

SCORE	0	1	2	3	4	5	6	7	8	9
STDEV	-1	0	1	2	3	4	5	6		

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.97

Times:	CPU	Total Elapsed
	00:01:25.97	00:01:36.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	4001

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.
Cut-off raised to 5.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

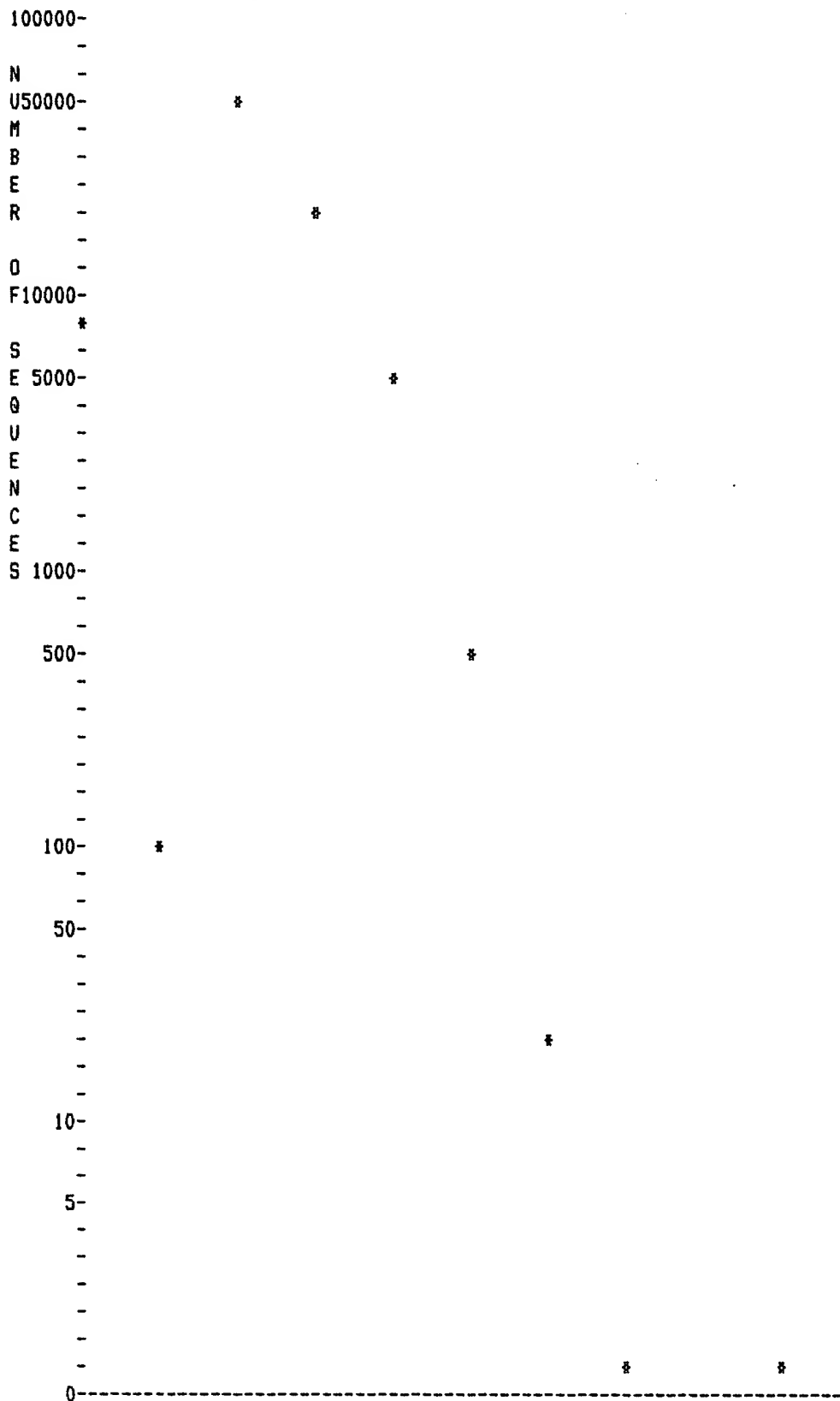
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. A42329	**** 7 standard deviations above mean **** autotaxin - human (fragments)	114	9	9	7.19	0
2. QQBEN4	**** 5 standard deviations above mean **** alkaline exonuclease (EC 3.1.	483	7	7	5.13	0
3. D35116	**** 4 standard deviations above mean **** hypothetical protein 2 (phnA	229	6	6	4.11	0
4. DGBEL5	uracil-DNA glycosylase (EC 3.	250	6	6	4.11	0
5. S32937	ureD protein - Klebsiella pne	270	6	6	4.11	0
6. A42887	urease - Klebsiella pneumonia	270	6	6	4.11	0
7. S02132	arginase (EC 3.5.3.1), hepatic	322	6	6	4.11	0
8. A26370	arginase (EC 3.5.3.1), hepatic	322	6	6	4.11	0
9. B26370	arginase (EC 3.5.3.1), hepatic	323	6	6	4.11	0
10. A28358	arginase (EC 3.5.3.1), hepatic	323	6	6	4.11	0
11. A26702	arginase (EC 3.5.3.1), hepatic	323	6	6	4.11	0
12. A31776	lactose permease - yeast (Klu	587	6	6	4.11	0
13. S28366	recombination repair protein	679	6	7	4.11	0
14. A48832	gpl38=cell surface glycoprote	730	6	6	4.11	0
15. KIRBF	6-phosphofructokinase (EC 2.7	749	6	6	4.11	0
16. A26550	6-phosphofructokinase (EC 2.7	780	6	6	4.11	0
17. JH0592	glutamate receptor chain KA-2	979	6	6	4.11	0
18. JH0589	glutamate receptor gamma-2 ch	979	6	6	4.11	0

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_5p.res made by on Thu 22 Sep 94 10:33:47-PDT.

Query sequence being compared:US-08-249-182-5 (1-10)
Number of sequences searched: 70848
Number of scores above cutoff: 4001

Results of the initial comparison of US-08-249-182-5 (1-10) with:
Data bank : PIR 41, all entries



CC Neutralising Domain (PND) of HIV-1. These peptides are used in
CC association with carriers which enhance immunogenicity. The carrier
CC may be the purified protein derivative (PPD) of tuberculin from
CC Mycobacterium tuberculosis. These peptide/carrier complexes can be
CC used as a vaccine for the treatment of HIV-infected and unaffected
CC individuals and/or for the transmission prevention of HIV-1. These
CC peptides may be linear or cyclic.

SQ Sequence 35 AA;
SQ 2 A; 5 R; 3 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 4 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
Residue Identity = 71% Matches = 5 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
CTRPNYNKRKRIHIGPGRFYTTKNIIGTIRGAHC
X X 10 20 30

15. US-08-249-182-5 (1-10)
R14336 HIV-1 amplifier peptide #20.

ID R14336 standard; Protein; 35 AA.
AC R14336;
DT 03-JAN-1992 (first entry)
DE HIV-1 amplifier peptide #20.
KW human immunodeficiency virus; vaccine; human retrovirus; AIDS;
KW acquired immunodeficiency syndrome; envelope glycoprotein.
OS Synthetic.
PN WD9114449-A.
PD 03-OCT-1991.
PF 16-MAR-1991; E00509.
PR 19-MAR-1990; US-494749.
PA (INSP) INST PASTEUR.
PI Girard M;
DR WPI; 91-310366/42.
PT Enhancing immunogenicity of envelope glyco:protein - for use as
PT vaccine or immuno:therapeutic drug especially against HIV, HTLV-I
PT and HTLV-II
PS Claim 12; Page 50; 71pp; English.
CC This peptide is one example of an HIV-1 amplifier peptide for use in
CC a composition for enhancing the immunogenicity of an envelope
CC glycoprotein of a virus. The sequence corresponds to a
CC neutralisation epitope and enhances the induction of persistent
CC neutralising antibodies in the host. The amplifier peptide is used
CC in addition to an envelope glycoprotein for priming the induction of
CC neutralising antibodies. The compositions are particularly
CC useful for vaccinating against HIV, SIV, HTLV-I and HTLV-II.
SQ Sequence 35 AA;
SQ 2 A; 5 R; 3 N; 1 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 3 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
Residue Identity = 71% Matches = 5 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
CTRPNYNKRKRIHIGPGRFYTTKNIIGDIRGAHC
X X 10 20 30

PD 02-DEC-1992.
 PF 29-MAY-1992; 109072.
 PR 31-MAY-1991; JP-129224.
 PA (KAGA) CHEMO SERO THERAPEUTIC RES INST.
 PI Eda Y, Osatomi K, Shiosaki K, Tokiyoshi S;
 DR WPI; 92-400517/49.
 PT Principal neutralising determinant peptide(s) of HIV gp120
 PT protein - used for diagnosing, preventing and treating HIV
 PT infection
 PS Example 1; Page 10; 26pp; English.
 CC DNA encoding HIV PND peptides was PCR amplified using genomic DNA
 CC from HIV-infected peripheral blood mononucleic cells as template.
 CC The amplified fragments were fused to beta-galactosidase coding
 CC sequence. E.coli transformants were cultured to produce the fusion
 CC protein. The expressed PND proteins were divided into groups based
 CC on their reactivity with neutralising antibodies and their amino
 CC acid sequence. The amino acid sequence was analysed using Robson's
 CC analytical program for protein secondary structure. Five groups
 CC were identified and 90% of all previously reported PND peptides
 CC were included in 3 main groups (i.e. Groups I, II and III).
 CC Group I PND peptides are those which have the structure XXXBBBX on
 CC the amino-terminal side of the GPGR motif (B = beta-strand
 CC structure and X = turn or coil structure). Vaccine preparations
 CC comprising representative peptides from each of the 5 groups can be
 CC used to develop vaccines able to recognise any HIV variant.
 CC See also Q31607-Q31608, R28996-R29000 and R29110-R29128.
 SQ Sequence 35 AA;
 SQ 2 A; 5 R; 3 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
 SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 4 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 71% Matches = 5 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 |||||
 CTRPNYNKRKRIHIGPGRFYTTKNIIGTIRQAHG
 X X 10 20 30

14. US-08-249-182-5 (1-10)
 R27857 gp120 PND fragment (m).

ID R27857 standard; peptide; 35 AA.
 AC R27857;
 DT 16-MAR-1993 (first entry)
 DE gp120 PND fragment (m).
 KW gp120; principal neutralising domain; PND; HIV-1; immunogenicity;
 KW purified protein derivative; PPD; tuberculin; vaccine; AIDS;
 KW Mycobacterium tuberculosis; transmission prevention.
 OS Human immunodeficiency virus type 1.
 PN W09217590-A.
 PD 15-OCT-1992.
 PF 02-APR-1992; E00735.
 PR 02-APR-1991; US-681624.
 PR 14-FEB-1992; US-837781.
 PA (INSS) SCHWEIZ SERUM & IMPFINST.
 PA (YESH) UNIV YESHIVA EINSTEIN COLLEGE.
 PI Bloom BR, Cryz S, Devash Y, Rubinstein A;
 DR WPI; 92-366260/44.
 PT Peptide conjugates of HIV gp120 principal neutralising domain -
 PT and carrier proteins as vaccines for treating and preventing HIV
 PT infection and/or transmission
 PS Claim 2; Page 54; 64pp; English.
 CC The sequences given in R27845-58 are fragments of the gp120 Principal

FT /note= "conserved PND motif"
 FT Region 8..14
 FT /note= "classification of PND peptides is
 FT determined by the predicted secondary
 FT structure of this region - see comments"
 PN EP-516135-A.
 PD 02-DEC-1992.
 PF 29-MAY-1992; 109072.
 PR 31-MAY-1991; JP-129224.
 PA (KAGA) CHEMO SERO THERAPEUTIC RES INST.
 PI Eda Y, Osatomi K, Shiosaki K, Tokiyoshi S;
 DR WPI; 92-400517/49.
 PT Principal neutralising determinant peptide(s) of HIV gp120
 PT protein - used for diagnosing, preventing and treating HIV
 PT infection
 PS Example 1; Page 10; 26pp; English.
 CC DNA encoding HIV PND peptides was PCR amplified using genomic DNA
 CC from HIV-infected peripheral blood mononucleic cells as template.
 CC The amplified fragments were fused to beta-galactosidase coding
 CC sequence. E.coli transformants were cultured to produce the fusion
 CC protein. The expressed PND proteins were divided into groups based
 CC on their reactivity with neutralising antibodies and their amino
 CC acid sequence. The amino acid sequence was analysed using Robson's
 CC analytical program for protein secondary structure. Five groups
 CC were identified and 90% of all previously reported PND peptides
 CC were included in 3 main groups (i.e. Groups I, II and III).
 CC Group I PND peptides are those which have the structure XXXBBBX on
 CC the amino-terminal side of the GPCR motif (B = beta-strand
 CC structure and X = turn or coil structure). Vaccine preparations
 CC comprising representative peptides from each of the 5 groups can be
 CC used to develop vaccines able to recognise any HIV variant.
 CC See Q31607-Q31608, R28995-R29000 and R29110-R29128.
 SQ Sequence 35 AA;
 SQ 2 A; 5 R; 3 N; 0 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
 SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 4 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 71% Matches = 5 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 ||||
 CTRPNYNKRRIHIGPGRFYTTKNIIGTIRQAHG
 X X 10 20 30

13. US-08-249-182-5 (1-10)
 R28995 Group I HIV gp120 PND peptide IIIMN.

ID R28995 standard; Protein; 35 AA.
 AC R28995;
 DT 02-APR-1993 (first entry)
 DE Group I HIV gp120 PND peptide IIIMN.
 KW Principal Neutralising determinant; Human immunodeficiency virus;
 KW vaccine; Robson's analytical method; polymerase chain reaction;
 KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.
 OS Human immunodeficiency virus.
 FH Key Location/Qualifiers
 FT Region 15..18
 FT /note= "conserved PND motif"
 FT Region 8..14
 FT /note= "classification of PND peptides is
 FT determined by the predicted secondary
 FT structure of this region - see comments"
 PN EP-516135-A.

PEEVTRPNYL

|||||
TRPNYNKRKRIHIGPGRAFYTTKNIIGDIRGAH
X X 10 20 30

11. US-08-249-182-5 (1-10)

R36586 Virus neutralising epitope of envelope glycoprotei

ID R36586 standard; peptide; 35 AA.
AC R36586;
DT 06-SEP-1993 (first entry)
DE Virus neutralising epitope of envelope glycoprotein of HIV.
KW Human immunodeficiency virus; gp120; gp160; EGP; VNE; immunity.
OS Synthetic.
PN W09308836-A.
PD 13-MAY-1993.
PF 28-OCT-1992; E02459.
PR 28-OCT-1991; US-782154.
PR 28-OCT-1991; US-782241.
PR 28-OCT-1991; US-782252.
PA (INSP) INST PASTEUR.
PI Girard M;
DR WPI; 93-167398/20.
PT Enhancing immunogenicity of viral envelope glycoprotein - by
PT co-administration of viral envelope glycoprotein itself, and an
PT oligopeptide derive.
PS Disclosure; Page 82; 107pp; English.
CC A novel method of enhancing the immunogenicity of an envelope
CC glycoprotein (EGP) of a virus (esp. HIV gp120 or gp160) in a host
CC comprises admin. to the host at least one EGP of the virus in an amt.
CC sufficient for priming vaccination and at least one peptide derived
CC from an amino acid sequence of the EGP (e.g. the sequence shown),
CC where the peptide comprises at least one virus-neutralisation
CC epitope (VNE). The complex is able to enhance the induction of
CC neutralising antibodies to the virus and to confer long lasting
CC immunity, longer than 6 months.
CC See also R36567-87.
SQ Sequence 35 AA;
SQ 2 A; 5 R; 3 N; 1 D; 0 B; 2 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 3 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
Residue Identity = 71% Matches = 5 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
CTRPYNKRKRIHIGPGRAFYTTKNIIGDIRGAHC
X X 10 20 30

12. US-08-249-182-5 (1-10)

R28996 Group I HIV gp120 PND peptide 8909C.

ID R28996 standard; Protein; 35 AA.
AC R28996;
DT 02-APR-1993 (first entry)
DE Group I HIV gp120 PND peptide 8909C.
KW Principal Neutralising determinant; Human immunodeficiency virus;
KW vaccine; Robson's analytical method; polymerase chain reaction;
KW Garnier-Osguthorpe-Robson method; GOR method; secondary structure.
OS Human immunodeficiency virus.
FH Key Location/Qualifiers
FT Region 15..18

FI breast and lung carcinomas; and melanoma
 PS Disclosure; Page 16; 17pp; English.
 CC This is the N-terminal amino acid sequence of a human tumour-
 CC associated cell surface glycoprotein antigen. The antigen is
 CC associated with non small cell lung carcinoma cells and also with
 CC other carcinomas including breast and colon carcinomas and melanomas.
 CC It can be used to produce a monoclonal antibody which can be used
 CC for both in vivo and in vitro diagnostic purposes such as the
 CC detection of malignant carcinomas.
 SQ Sequence 30 AA;
 SQ 1 A; 0 R; 1 N; 2 D; 0 B; 0 C; 3 Q; 0 E; 0 Z; 2 G; 0 H;
 SQ 0 I; 3 L; 0 K; 1 M; 1 F; 4 P; 1 S; 2 T; 0 W; 1 Y; 6 V;
 SQ 2 Others;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 || ||
 DVVVQAPTQVPGFLGDSVTLXPXYLQVPNMX
 10 X 20 X 30

10. US-08-249-182-5 (1-10)

R11210 Retroviral B-epitope containing peptide #20 of hyb

ID R11210 standard; Protein; 33 AA.
 AC R11210;
 DT 23-MAY-1991 (first entry)
 DE Retroviral B-epitope containing peptide #20 of hybrid molecule.
 KW retrovirus; env glycoprotein; B-epitope; immunodeficiency virus;
 KW HIV; SIV; AIDS.
 PN WD9102544-A.
 PD 07-MAR-1991.
 PF 17-AUG-1990; F00620.
 PR 18-AUG-1989; FR-011044.
 PA (INSP) INST PASTEUR.
 PA (UYCU-) UNIV CURIE P & M PARIS V.
 PI Girard M, Gluckman JC, Bahraoui EM;
 DR WPI; 91-087117/12.
 PT Vaccine compsns. which neutralise human immune deficiency virus -
 PT comprise a B epitope of retro-virus envelope glyco-protein and T
 PT epitope of distinct protein
 PS Claim 6; Page 36; 47pp; French.
 CC The peptide is a specific example of a B-epitope contg. peptide
 CC which can form a hybrid immunogenic molecule with a retroviral T-
 CC epitope. The B-epitope is chosen to be the major neutralisation
 CC epitope of the envelope glycoprotein of a pathogenic retrovirus.
 CC The T-epitope can be derived from a different protein of the
 CC same retrovirus or from the same protein from a different retrovirus.
 CC The hybrid molecule can also contain a minor epitope, especially a
 CC B-epitope from a conserved region of the HIV, SIV, HTLV-1 or
 CC HTLV-II env glycoprotein. The B-epitope-contg. peptide is joined to
 CC the T-epitope using eg tetanus toxin, KLH or HSA.
 CC See also R11191-R11209.
 SQ Sequence 33 AA;
 SQ 2 A; 5 R; 3 N; 1 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 3 G; 2 H;
 SQ 5 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 3 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

8. US-08-249-182-5 (1-10)

R40204 Sequence of peptide construct SP10MN(A).

ID R40204 standard; peptide; 24 AA.
AC R40204;
DT 05-FEB-1994 (first entry)
DE Sequence of peptide construct SP10MN(A).
KW Hybrid protein; synthetic protein; immunogenic peptide; tolerance;
KW synthetic toleragen.
OS Synthetic.
PN W09315750-A.
PD 19-AUG-1993.
PF 10-FEB-1993; U01207.
PR 10-FEB-1992; US-833429.
PA (HAYN/) HAYNES B F.
PI Haynes BF;
DR WPI; 93-272554/34.
PT Inducing immune tolerance to immunogenic peptide(s) or proteins -
PT by administering the peptide(s) or proteins coupled to a 2-20
PT aminoacid hydrophobic peptide
PS Example; Figure 10; 65pp; English.
CC The peptide is a variant of T1-SP10 peptide derived from HIVMN
CC envelope sequences. It comprises the following regions from the
CC envelope sequences: SP10 and A.
CC SP10MN(A) sequence is AAs 301-319 from HIVMN.
CC (A) sequence is AAs 32-324 from HIVMN and AAs 322-327 from HIVIIB.
CC A = additional HIV gp120 V3 loop sequences added to the original
CC synthetic peptide (SP10) sequence to add an additional neutralising
CC and CTL region to the HIV B cell determinant of the hybrid peptide.
SQ Sequence 24 AA;
SQ 1 A; 4 R; 2 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 2 G; 1 H;
SQ 2 I; 0 L; 3 K; 0 M; 1 F; 2 P; 0 S; 3 T; 0 W; 2 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
Residue Identity = 71% Matches = 5 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
CTRPNYNKRRIHIGPGRAFYTCK
X X 10 20

9. US-08-249-182-5 (1-10)

R15700 Human tumour-associated cell surface antigen N-ter

ID R15700 standard; peptide; 30 AA.
AC R15700;
DT 16-MAR-1992 (first entry)
DE Human tumour-associated cell surface antigen N-terminal.
KW Diagnosis; carcinoma; vaccine; therapy; lymphoma; glycoprotein.
OS Homo sapiens.
PN EP-460607-A.
PD 11-DEC-1991.
PF 04-JUN-1991; 109120.
PR 05-JUN-1990; US-533371.
PA (BRIM) BRISTOL-MYERS SQUIB.
PI Hellstrom I, Hellstrom KE, Marquardt H, Johnston J;
DR WPI; 91-363102/50.
PT Novel monoclonal antibody - to human tumour-associated cell
PT surface antigen useful in diagnosis and therapy of human colon,

CC R333347. While there was no detectable reactivity over background of
 CC MAb-01 with the peptides corresp. to AAs 302-316 or 322-336 of the
 CC V3 loop, binding of the antibody to the peptide representing AAs
 CC 3122-326 was apparent. The extent of this reactivity with other
 CC HIV-1 isolates was screened with peptides corresp. to the V3 loop
 CC region of HIV-1 isolates IIB, RF, CDC4, NY/5, Z6, Z2 and ELI
 CC (R33335-R33342). These results indicate that monoclonal antibody
 CC NM-01 recognizes an epitope of the V3 loop of gp120 of multiple
 CC HIV-1 isolates having the amino acid sequence R33343. NM-01 is also
 CC putatively reactive with the RF-like peptide set out in R33344.
 CC The variable region of the heavy and light chain of monoclonal
 CC antibody NM-01 were cloned by PCR and sequenced. Nucleotides 1-21
 CC and 334-363 of 837472 corresp. to the PCR primers used to amplify
 CC NM-01 light chain sequences and nucleotides 1-27 and 385-402 of
 CC 857471 corresp. to the PCR primers used to amplify NM-01 heavy chain
 CC sequences.

SQ Sequence 15 AA;
 SQ 0 A; 3 R; 2 N; 0 D; 0 B; 1 C; 0 Q; 0 E; 0 Z; 1 G; 1 H;
 SQ 2 I; 0 L; 2 K; 0 M; 0 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 71% Matches = 5 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 |||||
 CTRPNYNKRKRIHIG
 X X 10

7. US-08-249-182-5 (1-10)

P40385 Sequence of synthetic antigenic peptide 16 from gr

ID P40385 standard; peptide; 15 AA.
 AC P40385;
 DT 09-JAN-1992 (first entry)
 DE Sequence of synthetic antigenic peptide 16 from group C-src/yes
 DE family of oncoproteins.
 KW Vaccine; neoplasia; tumour location; diagnosis; oncogenic virus;
 KW antigen; oncoprotein; viral oncogene.
 PN W08403087-A.
 PD 16-AUG-1984.
 PF 14-FEB-1984; U00190.
 PR 14-FEB-1983; US-466329.
 PA (SENA/) SEN A.
 PI Sen A, Lerner RA, Houghten R, Bittle JL;
 DR WPI; 84-213376/34.
 PT Synthetic polypeptide(s) - useful for immunisation against
 PT neoplastic growth and in detection of neoplastic disease
 PS Example; Table 4, Page 53; 84pp; English.
 CC The synthetic peptides of the invention corresp. to an AA residue SQ
 CC of a first determinant domain of a first oncoprotein produced by
 CC cells transformed by an oncogenic virus. The determinant domain is
 CC vicinal to, but exclusive of, an active site of the oncoprotein.
 SQ Sequence 15 AA;
 SQ 0 A; 3 R; 0 N; 1 D; 0 B; 1 C; 0 Q; 3 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 1 F; 2 P; 0 S; 1 T; 1 W; 1 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 ||| ||

US *Simmondsia chinensis*.
 PN W09310241-A.
 PD 27-MAY-1993.
 PF 13-NOV-1992; U09863.
 PR 20-NOV-1991; US-796256.
 PR 21-AUG-1992; US-933411.
 PA (CALJ) CALGENE INC.
 PI Lardizabal KD, Lassner MW, Metz JG;
 DR WPI; 93-182556/22.
 PT Recombinant DNA construct for forming transgenic host cells which
 PT produce wax ester(s) - comprises nucleic acid sequence which
 PT encodes part of wax synthase protein and heterologous DNA
 PT sequence not naturally associated with wax synthase protein
 PT encoding sequence
 PS Disclosure; Page 46; 79pp; English.
 CC The sequences given in R37476-83 are jojoba wax synthase (fatty acyl:
 CC fatty alcohol acyltransferase) tryptic peptides. These peptides were
 CC used in the determination of the wax synthase amino acid sequence.
 CC The wax synthase and fatty acyl reductase coding sequences (see also
 CC 042838-39) were used in a construct in which the wax synthase gene is
 CC associated with a heterologous DNA sequence not naturally associated
 CC with it. This construct may be used to produce crop plants having a
 CC convenient source of wax esters. Wax esters can be used in a variety
 CC of industrial applications, including pharmaceuticals, cosmetics,
 CC detergents, plastics and lubricants. Production of wax esters in
 CC crop plants allows easier isolation than from traditional sources,
 CC eg. jojoba and sperm whale.
 SQ Sequence 10 AA;
 SQ 0 A; 0 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 3 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 0 L; 1 K; 0 M; 0 F; 1 P; 0 S; 2 T; 0 W; 1 Y; 2 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.15
 Residue Identity = 83% Matches = 5 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

```

      X   X 10
      PEEVTRPNYL
      |||||
  ETVVPEEVTK
      X   X 10
  
```

6. US-08-249-182-5 (1-10)

R33332 Sequence of peptide which corresp. to residues 302

ID R33332 standard; peptide; 15 AA.
 AC R33332;
 DT 06-JUL-1993 (first entry)
 DE Sequence of peptide which corresp. to residues 302-316 of the V3
 DE loop region of HIV-1MN gp120.
 KW Monoclonal antibody; NM-01; HIV-1; gp120; gp160.
 OS Synthetic.
 PN W09304090-A.
 PD 04-MAR-1993.
 PF 24-AUG-1992; U07111.
 PR 22-AUG-1991; US-748562.
 PA (NISP) NISSIN SHOKUHIN KAISHA LTD.
 PI Ohno T;
 DR WPI; 93-093943/11.
 PT Monoclonal antibodies against HIV-1 gp120 and gp160 proteins -
 PT for treating and preventing HIV-1 infection
 PS Example; Page 19; 57pp; English.
 CC Hybridoma cell line HB 10726 secretes MAb NM-01. In order to
 CC characterize the viral epitope recognized by NM-01, the Ab was
 CC screened reactivity with overlapping peptides corresp. to the amino
 CC acid sequence of the V3 loop region of HIV-1 gp120 (R33332, R33333,

ID R32359 standard; Protein; 980 AA.
 AC R32359;
 DT 17-JUN-1993 (first entry)
 DE Human KA-2 receptor.
 KW Kainate high affinity receptor; EAA2; excitatory amino acid family;
 KW assay; binding affinity; CNS disorders; drugs.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /note= "signal peptide"
 FT Peptide 19..962
 FT /note= "mature peptide"
 FT Domain 528..547
 FT /note= "transmembrane domain TM-1"
 FT Domain 572..590
 FT /note= "transmembrane domain TM-2"
 FT Domain 601..619
 FT /note= "transmembrane domain TM-3"
 FT Domain 786..806
 FT /note= "transmembrane domain TM-4"
 PN EP-529995-A.
 PD 03-MAR-1993.
 PF 25-AUG-1992; 307724.
 PR 27-AUG-1991; US-750081.
 PA (ALLE-) ALLELIX BIOPHARMACEUTICALS INC.
 PI Kamboj R, Nutt SL, Shekter L, Wosnick MA.
 DR WPI; 93-069002/09.
 DR N-PSDB; Q36930.
 PT Isolated polynucleotide encoding human excitatory amino acid-2
 PT receptor - useful for determining binding affinity of cpds. for
 PT the receptor in assaying for drugs to treat CNS disorders
 PS Claim 18; Fig 1; 29pp; English.
 CC The sequence is that of the human KA-2 receptor, a kainate high
 CC affinity receptor of the EAA2 (Excitatory Amino Acid) family. It
 CC can bind glutamate and has ligand-binding properties characteristic
 CC of kainate-type receptors.
 SQ Sequence 980 AA;
 SQ 79 A; 76 R; 36 N; 32 D; 0 B; 20 C; 30 Q; 67 E; 0 Z; 68 G; 18 H;
 SQ 55 I; 110 L; 28 K; 30 M; 38 F; 58 P; 81 S; 51 T; 13 W; 30 Y; 60 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.94
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                PEEVTRPNYL
                                ||| | ||
RDSQYETTDTCQILPKGVSVLGPSSSPASASTVSHICGEKEIPHIKVGPEETPRLLQYLRFASVSLYPSNE
      80      90      100      110      120 X      130      140

DVSLAVSRILKSFNYPASLCAKAECLLRLEELVRGF
      150      160      170      180

```

5. US-08-249-182-5 (1-10)

R37478 Jojoba wax synthase tryptic peptide, S01083.

ID R37478 standard; peptide; 10 AA.
 AC R37478;
 DT 20-SEP-1993 (first entry)
 DE Jojoba wax synthase tryptic peptide, S01083.
 KW Tryptic peptide; crop plant; fatty acyl:fatty alcohol acyltransferase;
 KW fatty acyl reductase; jojoba; wax ester; wax synthase; cosmetics;
 KW detergents; plastics; lubricants; pharmaceuticals; sperm whale.

CC the sequence is that encoded by the DNA insertion in plasmid pSV15/3
CC which may be used in the prodn. of attenuated non-primate herpes viruses.
CC These can be used as live vaccines and provide a safer vaccine than
CC currently available for e.g. pseudorabies virus of swine, infectious
CC bovine rhinotracheitis (IBR) virus or Marek's disease of fowl. This
CC is not the complete sequence as described in the specification but
CC is the only one given.

SQ Sequence 38 AA;
SQ 5 A; 3 R; 8 N; 4 D; 0 B; 0 C; 0 Q; 5 E; 0 Z; 1 G; 2 H;
SQ 0 I; 1 L; 0 K; 0 M; 0 F; 6 P; 0 S; 1 T; 0 W; 0 Y; 2 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.94
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
EAPGEEVTDPNANPNANPNANPNVDLERHHRADDR
X 10 X 20 30

3. US-08-249-182-5 (1-10)

R05306 Sequence of human arginase gene product.

ID R05306 standard; protein; 322 AA.
AC R05306;
DT 08-OCT-1990 (first entry)
DE Sequence of human arginase gene product.
KW Arginase; ds.
OS Homo sapiens.
PN J02117383-A.
PD 1-MAY-1990.
PF 26-OCT-1988; 268018.
PR 26-OCT-1988; JP-268018.
PA (TOYJ) Tosoh Corp.
DR WPI; 90-176227/23.
DR N-PSDB; 004714.
PT Prepn. of human arginase -
PT using transformed microorganism with vector having human
PT arginase coding DNA sequence.
PS Claim 3; Page 388-389; 25pp; Japanese.
CC A replicable vector expressing cDNA derived from the arginase gene
CC and coupled with a lactose promotor/operator system or a tryptophan
CC promotor/operator hybrid system allows production of the enzyme by
CC transformants.
CC See also 004715.
SQ Sequence 322 AA;
SQ 16 A; 13 R; 9 N; 19 D; 0 B; 3 C; 5 Q; 18 E; 0 Z; 37 G; 8 H;
SQ 23 I; 34 L; 24 K; 4 M; 9 F; 22 P; 21 S; 21 T; 2 W; 9 Y; 25 V;

Initial Score = 6 Optimized Score = 6 Significance = 3.94
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
PEEVTRPNYL
|||||
LDPSFTPATGTPVVGGLTYREGLYITGEIYKTGLLSGLDINEVNP SLGKTPEEVTRTVNTAVAILLACFGLA
240 250 260 270 280 X 290 X 300

REGNHKPIDYLNPPK
310 320

4. US-08-249-182-5 (1-10)

FH Key Location/Qualifiers
 FT Modified_site 6
 FT /note= "potentially glycosylated residue"
 PN US7822043-A.
 PD 01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 29. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 10 AA;
 SQ 0 A; 0 R; 1 N; 0 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 0 G; 0 H;
 SQ 0 I; 1 L; 0 K; 0 M; 0 F; 2 P; 0 S; 1 T; 0 W; 1 Y; 1 V;
 SQ 1 Others;

Initial Score = 9 Optimized Score = 9 Significance = 6.30
 Residue Identity = 90% Matches = 9 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 PEEVTRPNYL
 ||||| ||||
 PEEVTXPNYL
 X 10

2. US-08-249-182-5 (1-10)

R38708 Plasmid PSY1373 DNA insertion product.

ID R38708 standard; Protein; 38 AA.
 AC R38708;
 DT 25-NOV-1993 (first entry)
 DE Plasmid PSY1373 DNA insertion product.
 KW Attenuated; vaccine; herpes virus; non-primate; live; safer; IBR;
 KW infectious bovine rhinotracheitis; MDV; Marek's disease virus; fowl;
 KW pseudo-rabies; swine; infectious bursal disease virus.
 OS Synthetic.
 PN US5223424-A.
 PD 29-JUN-1993.
 PF 27-JUL-1988; 225032.
 PR 06-SEP-1985; US-773430.
 PR 27-JAN-1986; US-823102.
 PR 17-JUL-1986; US-887140.
 PR 02-SEP-1986; US-902887.
 PR 20-NOV-1986; US-933107.
 PR 27-JUL-1987; US-078519.
 PR 27-JUL-1988; US-225032.
 PA (PRUT-) PRUTECH RES & DEV.
 PI Chiang CH, Cochran MD, Macdonald RD.
 DR WPI; 93-219585/27.
 DR N-PSDB; 042766.
 PT Recombinant fusion proteins for vaccine - comprises antigenic
 PT sequences fused to viral sequences e.g. pseudo-rabies virus, used
 PT as vaccines
 PS Disclosure; Fig 51; 127pp; English.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 6 standard deviations above mean ****						
1. R37447	Autotaxin peptide ATX 29.	10	9	9	6.30	0
**** 3 standard deviations above mean ****						
2. R38708	Plasmid PSY1373 DNA insertion	38	6	6	3.94	0
3. R05306	Sequence of human arginase ge	322	6	6	3.94	0
4. R32359	Human KA-2 receptor.	980	6	6	3.94	0
5. R37478	Jojoba wax synthase tryptic p	10	5	5	3.15	0
6. R33332	Sequence of peptide which cor	15	5	5	3.15	0
7. P40385	Sequence of synthetic antigen	15	5	5	3.15	0
8. R40204	Sequence of peptide construct	24	5	5	3.15	0
9. R15700	Human tumour-associated cell	30	5	5	3.15	0
10. R11210	Retroviral B-epitope containi	33	5	5	3.15	0
11. R36586	Virus neutralising epitope of	35	5	5	3.15	0
12. R28996	Group I HIV gp120 PND peptide	35	5	5	3.15	0
13. R28995	Group I HIV gp120 PND peptide	35	5	5	3.15	0
14. R27857	gp120 PND fragment (n).	35	5	5	3.15	0
15. R14336	HIV-1 amplifier peptide #20.	35	5	5	3.15	0
16. R29128	Group V HIV gp120 PND peptide	36	5	5	3.15	0
17. R41926	CAP37 protein partial sequenc	40	5	5	3.15	0
18. R40201	Sequence of peptide construct	40	5	5	3.15	0
19. R40196	Sequence of peptide construct	40	5	5	3.15	0
20. R33217	HIV gp120 V3 loop immunogenic	40	5	5	3.15	0
21. R29228	Heteroconjugate antibody immu	40	5	5	3.15	0
22. R04461	Human immunodeficiency virus	40	5	5	3.15	0
23. R40202	Sequence of peptide construct	47	5	5	3.15	0
24. R41924	CAP37 protein coding sequence	50	5	5	3.15	0
25. P91880	Fusion protein coded for by o	50	5	6	3.15	0
26. R40203	Sequence of peptide construct	52	5	5	3.15	0
27. P92100	Fusion protein coded for by o	63	5	6	3.15	0
28. R31226	Fusion peptide contg. HIV-1 e	85	5	5	3.15	0
29. R42360	Papillomavirus E7 protein and	97	5	5	3.15	0
30. R26968	Human T lymphocyte receptor V	114	5	5	3.15	0
31. R28823	Alpha 6B integrin subunit cDN	141	5	5	3.15	0
32. R28824	Alpha 6A integrin subunit cDN	149	5	5	3.15	0
33. R28944	50 kD L-selectin ligand.	151	5	5	3.15	0
34. P50384	TNF analogue having modified	154	5	5	3.15	0
35. P50138	Rabbit tumor necrosis factor.	154	5	5	3.15	0
36. P50101	Sequence of a pure polypeptid	154	5	5	3.15	0
37. R05189	Tumoricidal polypeptide.	154	5	5	3.15	0
38. R05175	Tumoricidal polypeptide.	154	5	5	3.15	0
39. R03266	Rabbit tumour necrosis factor	154	5	5	3.15	0
40. P70557	Tumor necrosis factor.	154	5	5	3.15	0

1. US-08-249-182-5 (1-10)

R37447 Autotaxin peptide ATX 29.

ID R37447 standard; peptide; 10 AA.

AC R37447;

DT 22-JUL-1993 (first entry)

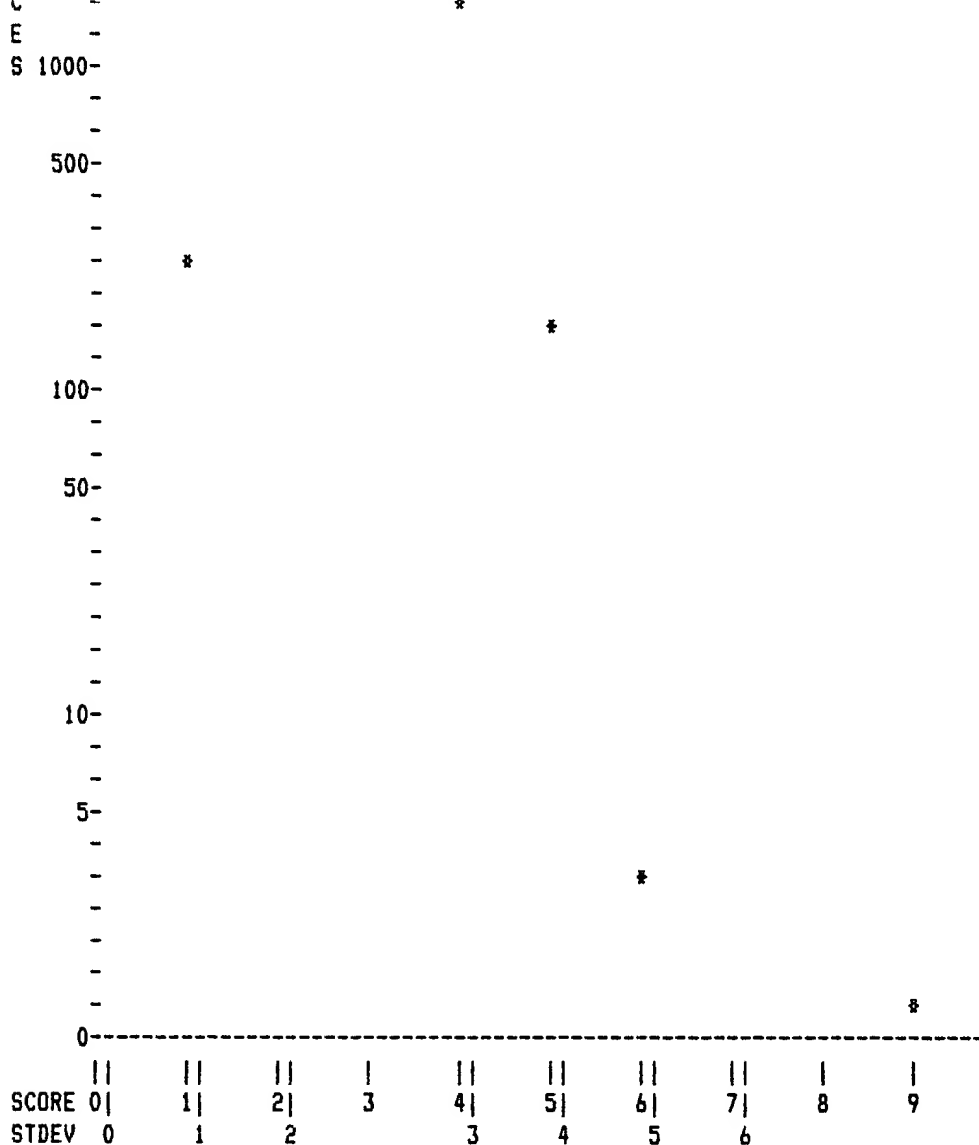
DE Autotaxin peptide ATX 29.

KW Cell motility stimulating; cancer metastasis; antibody; detection;

KW immunostains; disease outcome prediction; therapy choice;

KW cancer therapy; crosslinked toxins.

OS Synthetic.



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.27

Times:	CPU	Total Elapsed
	00:00:27.97	00:00:30.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	4616

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

15. US-08-249-182-4 (1-5)
VC18_BPML5 GENE 18 PROTEIN (GP18).

ID VC18_BPML5 STANDARD; PRT; 57 AA.

AC 005224;

DT 01-FEB-1994 (REL. 28, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)

DE GENE 18 PROTEIN (GP18).

CN 18.

OS MYCOBACTERIOPHAGE L5.

OC VIRIDAE; NOT YET CLASSIFIED.

RN []

RP SEQUENCE FROM N.A.

RM 93211282

RA HATFULL G.F., SARKIS G.J.;

RL MOL. MICROBIOL. 7:395-405(1993).

DR EMBL; Z18946; MLCGA.

SO SEQUENCE 57 AA; 6542 MW; 16653 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.56
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

VC18_BPML5
0AEVS
||||
MWSRKDHMRIGSTLNGFAEVSSEFAK0LIAIGWKVPRKPRNITKTATAPEEPKNEE
10 20 X 30 40 50

> 0 <
0110 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_5a.res made by on Thu 22 Sep 94 10:47:41-PDT.

Query sequence being compared:US-08-249-182-5 (1-10)
Number of sequences searched: 42145
Number of scores above cutoff: 4616

Results of the initial comparison of US-08-249-182-5 (1-10) with:
Data bank : A-Geneseg 15, all entries

100000-
N
-
US0000-
M
-
B
-
E
-
R
*
-
0
-
F10000-
S
-
E 5000-
Q
-
U
-
E
-
N

P1 NON_TER 42 42
SQ SEQUENCE 42 AA; 4727 MW; 10119 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.56
Residue Identity = 100% Matches = 4 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

                                X X
                                QAEVS
                                ||||
IDYRKCAQASQILKEHGMDKVIPLPELVCTMFHISGLSPQAEV
      10      20      30      40 X
```

12. US-08-249-182-4 (1-5)

THIO_EUBAC THIOREDOXIN (FRAGMENT).

ID THIO_EUBAC STANDARD; PRT; 45 AA.
AC P21610;
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
DE THIOREDOXIN (FRAGMENT).
OS EUBACTERIUM ACIDAMINOPHILUM.
OC PROKARYOTA; FIRMICUTES; IRREGULAR ASPOROGENOUS RODS; CORYNEFORM GROUP.
RN [1]
RP SEQUENCE.
RM 91139594
RA MEYER M., DIETRICH D., SCHMIDT B., ANDRESEN J.R.;
RL J. BACTERIOL. 173:1509-1513(1991).
CC -!- FUNCTION: THIOREDOXIN PARTICIPATES IN VARIOUS REDOX REACTIONS
CC THROUGH THE REVERSIBLE OXIDATION OF ITS ACTIVE CENTER DITHIOL,
CC TO A DISULFIDE, & CATALYZES DITHIOL-DISULFIDE EXCHANGE REACTIONS.
DR PROSITE; PS00194; THIOREDOXIN.
KW REDOX-ACTIVE CENTER; ELECTRON TRANSPORT.
FT DISULFID 32 35 REDOX-ACTIVE (BY SIMILARITY).
FT NON_TER 45 45
SQ SEQUENCE 45 AA; 5002 MW; 11785 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.56
Residue Identity = 80% Matches = 4 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

      X X
      QAEVS
      ||||
MALLVEIDKDGQAEVLEAEGYVLVDYFSDGCVPCALMPDVXIL
      10 X X 20      30      40
```

13. US-08-249-182-4 (1-5)

PSBK_CHLRE PHOTOSYSTEM II 4 KD REACTION CENTRE PROTEIN PRECUR

ID PSBK_CHLRE STANDARD; PRT; 46 AA.
AC P18263;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE PHOTOSYSTEM II 4 KD REACTION CENTRE PROTEIN PRECURSOR.
GN PSBK.
OS CHLAMYDOMONAS REINHARDTII.
OG CHLOROPLAST.
OC EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE); VOLVOCALES;
OC CHLAMYDOMONADACEAE.
RN [1]
RP SEQUENCE FROM N.A.

AC STRAIN-CC123;
 RM 90370493
 RA SILK G.W., DE LA CRUZ F., WU M.;
 RL NUCLEIC ACIDS RES. 18:4930-4930(1990).
 CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER
 CC OF PHOTOSYSTEM II.
 DR EMBL; X53413; CHCRPSBK.
 DR PIR; S11162; S11162.
 KW CHLOROPLAST; PHOTOSYSTEM II; SIGNAL.
 FT SIGNAL 1 9 POTENTIAL.
 FT CHAIN 10 46 4 KD REACTION CENTER PROTEIN.
 SQ SEQUENCE 46 AA; 5060 MW; 12441 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.56
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 || ||
 MTTLALVLAKLPEAYAPFAPIVDLPVIPVFFILLAFVWQAASVFR
 10 20 30 40 X

14. US-08-249-182-4 (1-5)

PSBK_MARPO PHOTOSYSTEM II 4 KD REACTION CENTRE PROTEIN PRECUR

ID PSBK_MARPO STANDARD; PRT; 55 AA.
 AC P10348;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PHOTOSYSTEM II 4 KD REACTION CENTRE PROTEIN PRECURSOR.
 GN PSBK.
 OS MARCHANTIA POLYMORPHA (LIVERWORT).
 OG CHLOROPLAST.
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; BRYOPHYTA; HEPATICOPSIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA OHYAMA K.;
 RL SUBMITTED (OCT-1986) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP COMPLETE GENOME.
 RA OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SAND T., SAND S.,
 RA UMESONO K., SHIKI Y., TAKEUCHI M., CHANG Z., AOTA S., INOKUCHI H.,
 RA OZEKI H.;
 RL NATURE 322:572-574(1986).
 CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER
 CC OF PHOTOSYSTEM II.
 DR EMBL; X04465; CHMPXX.
 DR PIR; A05024; A05024.
 DR PIR; S01585; S01585.
 KW CHLOROPLAST; PHOTOSYSTEM II; SIGNAL.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 55 4 KD REACTION CENTER PROTEIN.
 SQ SEQUENCE 55 AA; 6442 MW; 18114 CN;

Initial Score = 4 Optimized Score = 4 Significance = 3.56
 Residue Identity = 80% Matches = 4 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X X
 QAEVS
 || ||
 MFNIYLENAFYLNIGITFAKLPEAYSIFDPIVDVMPPIPLFFFLAFVWQASVSFR
 10 20 30 40 50 X

50 21 A; 12 R; 29 N; 19 D; 0 B; 8 C; 16 Q; 28 E; 0 Z; 18 G; 8 H;
50 31 I; 34 L; 37 K; 6 M; 18 F; 22 P; 35 S; 29 T; 12 W; 17 Y; 36 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPETTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
      260      270      280
```

3. US-08-249-182-7 (1-10)

R14487 Soluble interferon-alpha/beta receptor.

ID R14487 standard; Protein; 436 AA.
AC R14487;
DT 16-JAN-1992 (first entry)
DE Soluble interferon-alpha/beta receptor.
KW IFN; autoimmune disease; graft rejection; histocompatibility.
OS Homo sapiens.
PN FR2657881-A.
PD 09-AUG-1991.
PF 05-FEB-1990; 001298.
PR 05-FEB-1990; FR-001298.
PA (EUBI-) LAB EURO BIOTECHNO.
PI Eid P, Gresser I, Lutfalla G, Meyer F, Mogensen KE;
PI Tovey MG, Uze G;
DR WPI; 91-319778/44.
DR N-PSDB; Q14239.
PT New water-soluble polypeptide(s) with affinity for IFN-alpha and
PT beta - used to treat e.g. lupus erythematosus, Behcet's disease,
PT aplastic anaemia, diabetes mellitus, rheumatoid arthritis, etc.
PS Claim 2; Page 45; 52pp; French.
CC The transmembrane and cytoplasmic domains of the native IFN receptor
CC have been deleted to obtain a soluble, circulating form of the
CC receptor. Potentially immunogenic epitopes have thus been eliminated.
CC Derivatives obtained by substitution or deletion of this sequence
CC are also claimed as are hybrid molecules comprising the soluble
CC receptor (or deriv.) and an immunoglobulin such as IgG1.
CC See also Q14240.
SQ Sequence 436 AA;
SQ 21 A; 12 R; 29 N; 19 D; 0 B; 8 C; 16 Q; 28 E; 0 Z; 18 G; 8 H;
SQ 31 I; 34 L; 37 K; 6 M; 18 F; 22 P; 35 S; 29 T; 12 W; 17 Y; 36 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPETTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
      260      270      280
```

4. US-08-249-182-7 (1-10)

R29583 Human activin receptor.

PU 01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 48. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 10 AA;
 SQ 0 A; 0 R; 2 N; 0 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 2 P; 1 S; 0 T; 0 W; 1 Y; 0 V;

Initial Score = 8 Optimized Score = 8 Significance = 5.90
 Residue Identity = 80% Matches = 8 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
| | | | | | | |
SPPFENINLY
      X      10
  
```

2. US-08-249-182-7 (1-10)

R28495 Sequence of a soluble form of the interferon (IFN)

ID R28495 standard; Protein; 436 AA.
 AC R28495;
 DT 31-MAR-1993 (first entry)
 DE Sequence of a soluble form of the interferon (IFN) receptor
 DE with a high affinity for IFN-alpha and -beta.
 KW Interferon receptor; alpha-interferon; beta-interferon.
 OS Synthetic.
 PN W09218626-A.
 PD 29-OCT-1992.
 PF 17-APR-1991; F00318.
 PR 17-APR-1991; WO-F00318.
 PA (EUBI-) LAB EURO BIOTECHNOLOGIE.
 PI Eid P, Gresser I, Lutfalla G, Meyer F, Mogensen KE,
 PI Tovey M, Uze G;
 DR WPI; 92-382110/46.
 DR N-PSDB; Q30532.
 PT Water soluble polypeptide(s) strongly bind interferon(s) alpha
 PT and beta - useful as immunosuppressants, for treating auto:immune
 PT diseases and transplant rejection
 PS Claim 2; Fig 1; 58pp; English.
 CC DNA encoding the water-soluble polypeptide with a high affinity for
 CC IFN-alpha and -beta is isolated by PCR, using appropriate
 CC oligonucleotides as primers and cloned cDNA as template. For example,
 CC bacteriophage lambda ZAP, containing the entire coding sequence of
 CC the IFN-alpha and -beta receptor (Q30533), was incubated with oligos
 CC Q30534 and Q30535. R28496 represents the complete receptor. R28495
 CC lacks the transmembrane and cytoplasmic domains. Both forms bind
 CC IFN in the same way as antibodies so are immunosuppressants e.g. for
 CC treating autoimmune diseases and graft rejection. They lack the
 CC toxic side-effects of known immunosuppressants such as steroids.
 SQ Sequence 436 AA;

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Init. Opt.		Sig.	Frame
		Length	Score		
**** 5 standard deviations above mean ****					
1. R37449	Autotaxin peptide ATX 48.	10	8	8	5.90 0
**** 4 standard deviations above mean ****					
2. R28495	Sequence of a soluble form of	436	6	6	4.21 0
3. R14487	Soluble interferon-alpha/beta	436	6	6	4.21 0
4. R29583	Human activin receptor.	513	6	6	4.21 0
5. R29581	Mouse activin receptor.	513	6	6	4.21 0
6. R42635	Human interferon receptor.	557	6	6	4.21 0
7. R28496	Sequence of a soluble form of	557	6	6	4.21 0
8. R14488	Complete interferon-alpha/bet	557	6	6	4.21 0
9. R11958	Human alpha-interferon recept	557	6	6	4.21 0
**** 3 standard deviations above mean ****					
10. R10607	Peptide with motilin-like act	13	5	5	3.37 0
11. R10606	Peptide with motilin-like act	14	5	5	3.37 0
12. R10605	Peptide with motilin-like act	15	5	5	3.37 0
13. R10601	Peptide with motilin-like act	15	5	5	3.37 0
14. R10604	Peptide with motilin-like act	17	5	5	3.37 0
15. R10603	Peptide with motilin-like act	18	5	5	3.37 0
16. R10602	Peptide with motilin-like act	18	5	5	3.37 0
17. R10600	Peptide with motilin-like act	18	5	5	3.37 0
18. P81032	Sequence (Formula 1) of a 13-	22	5	5	3.37 0
19. R03267	Motilin-like polypeptide.	22	5	5	3.37 0
20. P82050	13-Leu motilin	22	5	5	3.37 0
21. R04201	N-terminal of deacetoxycephal	22	5	5	3.37 0
22. P92007	Polypeptide with motilin-like	22	5	5	3.37 0
23. R41237	Motilin-like polypeptide #3.	23	5	5	3.37 0
24. R41235	Motilin-like polypeptide #1.	23	5	5	3.37 0
25. R41236	Motilin-like polypeptide #2.	23	5	5	3.37 0
26. R41194	Motilin-like polypeptide gene	23	5	5	3.37 0
27. R05893	Motilin-like peptide.	23	5	5	3.37 0
28. P92005	Polypeptide with motilin-like	23	5	5	3.37 0
29. P92006	Polypeptide with motilin-like	23	5	5	3.37 0
30. R05892	Motilin-like peptide.	25	5	5	3.37 0
31. P83215	Sequence of a 13-Leu analogue	26	5	5	3.37 0
32. R07040	Motilin-like peptide.	28	5	5	3.37 0
33. R07039	Motilin-like peptide.	29	5	5	3.37 0
34. R12911	Fragile X syndrome related pe	36	5	5	3.37 0
35. R05364	[Leu13] motilin with leader p	48	5	5	3.37 0
36. P83216	Sequence of dimer of a 13-Leu	52	5	5	3.37 0
37. R29163	PRP2.	114	5	5	3.37 0
38. R10981	Motilin/white salmon growth h	114	5	5	3.37 0
39. P80459	Sequence corresp. to the nucl	115	5	5	3.37 0
40. P91380	Porcine prepronotilin.	119	5	5	3.37 0

1. US-08-249-182-7 (1-10)

R37449 Autotaxin peptide ATX 48.

ID R37449 standard; peptide; 10 AA.

AC R37449;

DT 22-JUL-1993 (first entry)

DE Autotaxin peptide ATX 48.

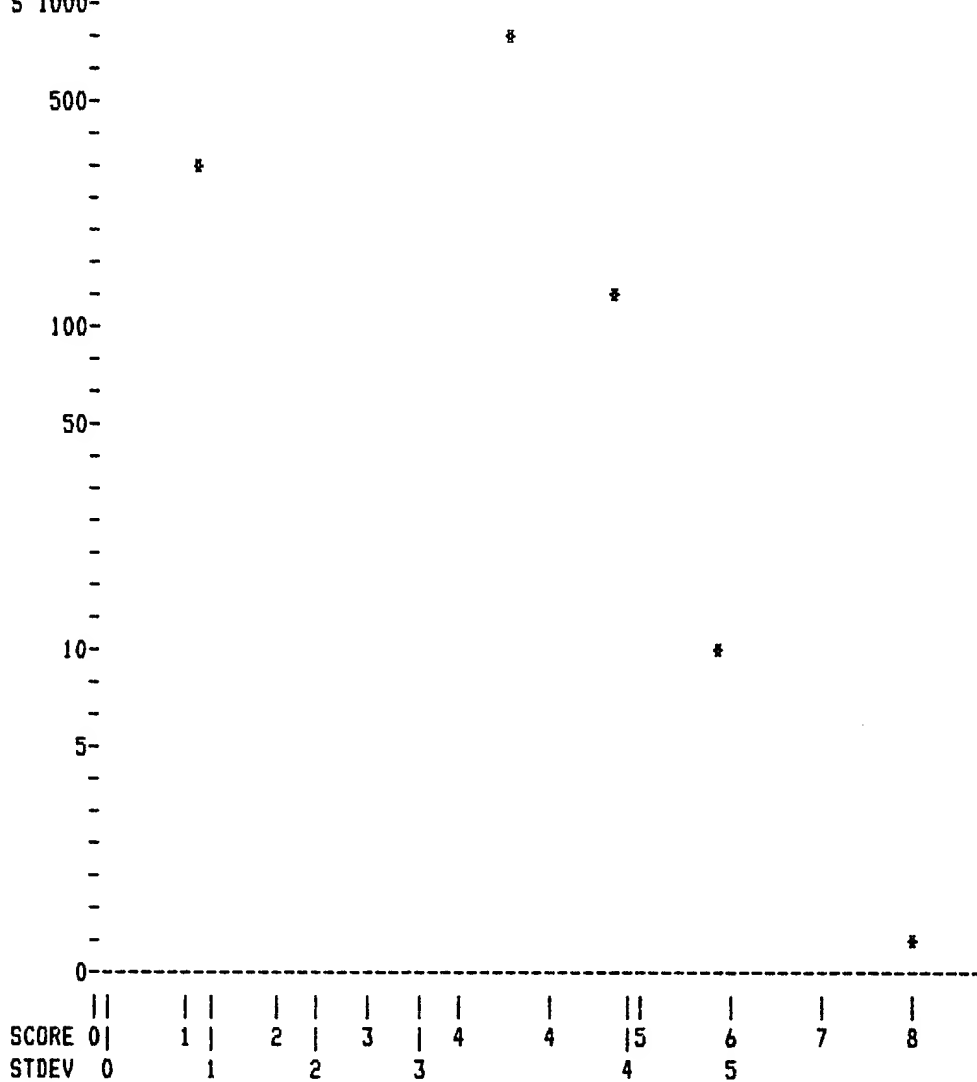
KW Cell motility stimulating; cancer metastasis; antibody; detection;

KW immunostains; disease outcome prediction; therapy choice;

KW cancer therapy; crosslinked toxins.

OS Synthetic.

PN US7822043-A.



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.19
Times:	CPU	Total Elapsed	
	00:00:27.94	00:00:30.00	
Number of residues:	5287517		
Number of sequences searched:	42145		
Number of scores above cutoff:	4126		

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

DE CYSTATIN; CULUSIROM (THIOL PROTEINASE INHIBITOR).
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 QC EUTHERIA; ARTIODACTYLA.
 RN [1]
 RP SEQUENCE.
 RM 85231205
 RA HIRADO M., TSUNASAWA S., SAKIYAMA F., NIINOBE M., FUJII S.;
 RL FEBS LETT. 186:41-45(1985).
 CC -!- FUNCTION: THIS IS A THIOL PROTEINASE INHIBITOR.
 CC -!- SIMILARITY: THIS IS A TYPE 2 CYSTATIN.
 DR PIR; A01271; UDBO.
 DR PRDSITE; PS00287; CYSTATIN.
 KW THIOL PROTEASE INHIBITOR.
 FT ACT_SITE 4 4 REACTIVE SITE.
 FT SITE 48 52 SECONDARY AREA OF CONTACT.
 FT DISULFID 66 76 BY SIMILARITY.
 FT DISULFID 90 110 BY SIMILARITY.
 SQ SEQUENCE 112 AA; 12789 MW; 68828 CN;

Initial Score = 5 Optimized Score = 6 Significance = 3.52
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 I III II
 ARKQVVSGMNYFLDELGRITCTKSQANLDSCPFHNQPHLKREKLCSEFQVVVVPWMNTINLVKFSCQD
 50 60 70 80 90 X 100 X 110

> 0 <
 0| |0 IntelliGenetics
 > 0 <

Seq. 7

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_7a.res made by on Thu 22 Sep 94 10:56:41-PDT.

Query sequence being compared:US-08-249-182-7 (1-10)
 Number of sequences searched: 42145
 Number of scores above cutoff: 4126

Results of the initial comparison of US-08-249-182-7 (1-10) with:
 Data bank : A-GeneSeq 15, all entries

100000-
 N -
 U50000-
 M -
 B -
 E *
 R - *
 D -
 F10000-
 S -
 E 5000-
 Q - *
 U -
 E -
 N -
 C -
 E -

Initial Score = 5 Optimized Score = 5 Significance = 3.52
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```
      X      X
      YDVPWNETI
      |||  ||
AYKTVLKTPSGEFTLDVPEGTTILDAAEEAGYDLPFSCRAGACSSCLGKVVSGSVDSSEGSFLDDGQMEEGF
      10   X   20   X   30       40       50       60       70
```

V

14. US-08-249-182-6 (1-9)
FER1_EQUAR FERREDOXIN I.

ID FER1_EQUAR STANDARD; PRT; 95 AA.
AC P00235;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE FERREDOXIN I.
OS EQUISETUM ARVENSE (FIELD HORSETAIL) (COMMON HORSETAIL).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; EQUISETOPHYTA; SPHENOPSIDA;
OC EQUISETALES; EQUISETACEAE.
RN [1]
RP SEQUENCE.
RM 77249492
RA HASE T., WADA K., MATSUBARA H.;
RL J. BIOCHEM. 82:277-286(1977).
CC -!- FUNCTION: FERREDOXIN ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- THIS FERREDOXIN BINDS A SINGLE 2FE-2S CLUSTER.
DR PIR; A04609; FEEQ1F.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN.
KW ELECTRON TRANSPORT; IRON-SULFUR; CHLOROPLAST.
FT METAL 38 38 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 43 43 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 76 76 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10098 MW; 47218 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.52
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```
      X      X
      YDVPWNETI
      |||  ||
AYKTVLKTPSGEFTLDVPEGTTILDAAEEAGYDLPFSCRAGACSSCLGKVVSGSVDSSEGSFLDDGQMEEGF
      10   X   20   X   30       40       50       60       70
```

V

15. US-08-249-182-6 (1-9)
CYTC_BOVIN CYSTATIN, COLOSTRUM (THIOL PROTEINASE INHIBITOR).

ID CYTC_BOVIN STANDARD; PRT; 112 AA.
AC P01035;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)

12. US-08-249-182-6 (1-9)
VKIL_BPP22 KIL PROTEIN.

ID VKIL_BPP22 STANDARD; PRT; 62 AA.
AC P14111;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE KIL PROTEIN.
GN KIL.
OS BACTERIOPHAGE P22.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 89293845
RA SEMERJIAN A.V., MALLOY D.C., POTEETE A.R.;
RL J. MOL. BIOL. 207:1-13(1989).
CC -!- FUNCTION: P22 KIL IS ESSENTIAL FOR LYTIC GROWTH IN THE ABSENCE OF
CC ABC. EXPRESSION OF P22 KIL CAUSES FILAMENTATION AND CELL DEATH.
DR EMBL; X15637; POP22PL.
DR PIR; S04248; VDBP22.
SQ SEQUENCE 62 AA; 6950 MW; 15686 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.52
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

X      X
YDVPWNETI
  || ||
MTIVPVNGTILVQGNREFNKL YEASFPDTKEGNSAAYAWASSIAMGWEDCQDEDWNRNH
X      10      20      30      40      50      60
```

13. US-08-249-182-6 (1-9)
FER1_EQUTE FERREDOXIN I.

ID FER1_EQUTE STANDARD; PRT; 95 AA.
AC P00234;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE FERREDOXIN I.
OS EQUISETUM TELMATEIA (GIANT HORSETAIL).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; EQUISETOPHYTA; SPHENOPSIDA;
OC EQUISETALES; EQUISETACEAE.
RN [1]
RP SEQUENCE.
RM 77249491
RA HASE T., WADA K., MATSUBARA H.;
RL J. BIOCHEM. 82:267-276(1977).
CC -!- FUNCTION: FERREDOXIN ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- THIS FERREDOXIN BINDS A SINGLE 2FE-2S CLUSTER.
DR PIR; A00240; FEE01.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN.
KW ELECTRON TRANSPORT; IRON-SULFUR; CHLOROPLAST.
FT METAL 38 38 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 43 43 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 76 76 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10097 MW; 47161 CN;

DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 375.7 KD PROTEIN ZK112.7 IN CHROMOSOME III.
 GN ZK112.7.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; ACCELEMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA DU Z.;
 RL SUBMITTED (MAY-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- SIMILARITY: TO D.MELANOGASTER FAT TUMOR SUPPRESSOR.
 DR EMBL; L14324; CEZK112.
 DR WORMPEP; ZK112.7; CE00378.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 3343 AA; 375745 MW; 22105723 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 || |||
 VTCIQKNSTCQPTLVGDSASRLVSRSSSVIFDLPLKKLTARCFCSGIDCYDDTTNETIQKTQKINVITTC
 2930 2940 2950 2960 2970 2980 2990
 DIDCGPRGKCFMEESSQPICRGGGFESMYSCERADD
 3000 3010 3020

11. US-08-249-182-6 (1-9)
 TES1_RAT TESTIN 1 (CMB-22) (FRAGMENT).

ID TES1_RAT STANDARD; PRT; 30 AA.
 AC P15242;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
 DE TESTIN 1 (CMB-22) (FRAGMENT).
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SERTOLI CELLS;
 RM 90078247
 RA CHENG C.Y., GRIMA J., STAHLER M.S., LOCKSHIN R.A.;
 RL J. BIOL. CHEM. 264:21386-21393(1989).
 CC -!- FUNCTION: NOT KNOWN.
 CC -!- SUBCELLULAR LOCATION: SERTOLI CELLS SECRETORY PROTEIN.
 DR PIR; A34199; A34199.
 KW TESTIS.
 FT NON_TER 30 30
 SQ SEQUENCE 30 AA; 3731 MW; 4350 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.52
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 || |||
 TPDPSLDVEWNEWRTKHGKTYNMNEERLKR
 X 10 X 20 30

TSLSNEGRLVGLFASKSPFADQVNDGLKNTISNRATAESDLVFLGLIGITDPPRNETAGAVKRPQAGLINV
600 610 620 630 640 650 X 660

HMLTGDFVGTAKAIAQEVGILPTNLYHVSQEIIVDSMV
670 680 690 700

9. US-08-249-182-6 (1-9)

CYGS_STRPU SPERACT RECEPTOR PRECURSOR (GUANYLATE CYCLASE) (EC

ID CYGS_STRPU STANDARD; PRT: 1125 AA.
AC P16065;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE SPERACT RECEPTOR PRECURSOR (GUANYLATE CYCLASE) (EC 4.6.1.2).
DS STRONGYLOCENTROTUS PURPURATUS (PURPLE SEA URCHIN).
OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA;
OC EUECHINOIDEA.
RN [1]
RP SEQUENCE FROM N.A.
RM 89197965
RA THORPE D.S., GARBERS D.L.;
RL J. BIOL. CHEM. 264:6545-6549(1989).
CC -!- FUNCTION: IMPLICATED AS A CELL-SURFACE RECEPTOR ON SPERMATOZOA
CC FOR 'SPERACT' A CHEMOTACTIC PEPTIDE, AND ON VARIOUS OTHER CELLS
CC AS A RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE.
CC -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: TO OTHER GUANYLATE CYCLASES.
CC -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC DOMAIN OF PROTEIN KINASES.
DR EMBL; M22444; SPGUA.
DR PIR; A33535; OYURCP.
DR PIR; A30856; A30856.
DR PROSITE; PS00452; GUANYLATE_CYCLASES.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; PHOSPHORYLATION; LYASE;
KW CGMP SYNTHESIS; SIGNAL.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 1125 GUANYLATE CYCLASE.
FT DOMAIN 22 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 POTENTIAL.
FT DOMAIN 532 1125 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 185 185 POTENTIAL.
FT CARBOHYD 409 409 POTENTIAL.
SQ SEQUENCE 1125 AA; 126256 MW; 6471318 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
||| |||
IAVLPETFEMVSIFFSIDVGFALTSAASTPIQVVNLLNDLYTLFDAIISNYDVYKVETIGDAYMLVSGPLPLR
910 920 930 940 950 X 960 970

NGDRHAGQIASTAHLLSVKGFIVPHKPEVFLKLRI
980 990 1000 1010

10. US-08-249-182-6 (1-9)

YOG7_CAEEL HYPOTHETICAL 375.7 KD PROTEIN ZK112.7 IN CHROMOSOM

ID YOG7_CAEEL STANDARD; PRT: 3343 AA.
AC P34616;


```

Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
||| |||
YSLSNEGLRVLGFAKSFYKQVNDQDLKNITSNRATAESDLVFLGLIGIYDPPRNETAGAVKKFHGAGINV
600 610 620 630 640 650 X 660

HMLTGDFVGTAKAIAQEVGILPTNLVHYSQEIYDSMV
670 680 690 700

```

8. US-08-249-182-6 (1-9)

ATN1_YEAST SODIUM TRANSPORT ATPASE 1 (EC 3.6.1.-).

```

ID ATN1_YEAST STANDARD; PRT; 1091 AA.
AC P13587;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE SODIUM TRANSPORT ATPASE 1 (EC 3.6.1.-).
GN ENA1 OR PMR2.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RM 89324047
RA RUDOLPH H.K., ANTEBI A., FINK G.R., BUCKLEY C.M., DORMAN T.E.,
RA LEVITRE J., DAVIDOW L.S., MAO J.-I., MOIR D.T.;
RL CELL 58:133-145(1989).
RN [2]
RP SEQUENCE OF 534-1091 FROM N.A.
RC STRAIN=7305B;
RM 91260670
RA MARTINEZ R., LATREILLE M.-T., MIRANDE M.;
RL MOL. GEN. GENET. 227:149-154(1991).
RN [3]
RP PRELIMINARY SEQUENCE OF 534-1091 FROM N.A.
RC STRAIN=X2180;
RM 89054027
RA MIRANDE M., WALLER J.-P.;
RL J. BIOL. CHEM. 263:18443-18451(1988).
CC -!- FUNCTION: THIS MAGNESIUM-DEPENDENT, ENZYME PROBABLY CATALYZES THE
CC HYDROLYSIS OF ATP COUPLED WITH THE TRANSPORT OF THE CALCIUM.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES).
DR EMBL; M25489; SCPMR2.
DR EMBL; X58626; SCPMR2G.
DR EMBL; J04186; SCKRS1A.
DR PIR; S05788; PWBVR2.
DR PROSITE; PS00154; ATPASE_E1_E2.
KW HYDROLASE; SODIUM TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;
KW MAGNESIUM; ATP-BINDING; MULTIGENE FAMILY.
FT MOD_RES 369 369 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 1091 AA; 120357 MW; 6231113 CN;

```

```

Initial Score = 6 Optimized Score = 6 Significance = 4.69
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

```

X X
YDVPWNETI
||| |||

```

DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE AL1 PROTEIN (ORF C1).
 OS PANICUM STREAK VIRUS.
 DC VIRIDAE; SS-DNA NONENVELOPED VIRUSES; GEMINIVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92268861
 RA BRIDDON R.W., LUNNESS P., CHAMBERLAIN L.C., BRUNDISH H.,
 RA PINNER M.S., MARKHAM P.G.;
 RL J. GEN. VIROL. 73:1041-1047(1992).
 CC -!- SIMILARITY: TO AL1 PROTEIN IN OTHER GEMINIVIRUSES.
 DR EMBL; X60168; PSGIITDNA.
 DR PIR; J01552; J01552.
 SQ SEQUENCE 323 AA; 37026 MW; 566018 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||  |||
HATSREEYLSLVQSSLPYDWATKLNIFYEYSASRLFPDIAEPYTNPHPTTEYDLHCNETIEDWLKPNYQVSP
    160      170      180      190      200 X      210      220

QAYKLLPEPSCLSLEGAIADLEWLDDTTRMLQEKEREA
    230      240      250      260
  
```

7. US-08-249-182-6 (1-9)

ATN2_YEAST SODIUM TRANSPORT ATPASE 2 (EC 3.6.1.-).

ID ATN2_YEAST STANDARD; PRT; 1091 AA.
 AC 001896;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE SODIUM TRANSPORT ATPASE 2 (EC 3.6.1.-).
 GN ENA2.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 DC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBY746;
 RM 93173113
 RA GARCIABLAS B., RUBIO F., QUINTERO F.J., BANUELOS M.A., HARD R.,
 RA RODRIGUEZ-NAVARRO A.;
 RL MOL. GEN. GENET. 236:363-368(1993).
 CC -!- FUNCTION: THIS MAGNESIUM-DEPENDENT, ENZYME PROBABLY CATALYZES THE
 CC HYDROLYSIS OF ATP COUPLED WITH THE TRANSPORT OF SODIUM AND
 CC LITHIUM.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES).
 DR EMBL; X67136; SCENA2.
 DR PIR; S25007; S25007.
 DR PROSITE; PS00154; ATPASE_E1_E2.
 KW HYDROLASE; SODIUM TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;
 KW MAGNESIUM; ATP-BINDING; MULTIGENE FAMILY.
 FT MOD_RES 369 369 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 1091 AA; 120317 MW; 6218101 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69

Initial Score = 6 Optimized Score = 6 Significance = 4.69
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| |||
EADPKAYAGHVFRSFDANS DGTLD FKEYVIALHMTSAGKTN QKLEWAFSLYD V DNGTISKNEVLEIVTAIF
  60      70      80      90      100     110     X 120

```

```

KMISPEDTKHLPEDENTPEKRAEKI WGF GKKDDDKL
130      140      150      160

```

5. US-08-249-182-6 (1-9)

FLGL_ECOLI FLAGELLAR HOOK-ASSOCIATED PROTEIN 3 (HAP3) (HOOK-F

ID FLGL_ECOLI STANDARD; PRT; 317 AA.
 AC P29744;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE FLAGELLAR HOOK-ASSOCIATED PROTEIN 3 (HAP3) (HOOK-FILAMENT JUNCTION
 DE PROTEIN).
 GN FLGL OR FLAT OR FLAU.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / CS520;
 RA FAHRNER K.A., BLOCK S.M., KRISHNASWAMY S., PARKINSON J.S., BERG H.C.;
 RL SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE OF 292-317 FROM N.A.
 RC STRAIN=K12;
 RM 93078265
 RA CASAREGOLA S., JACQ A., LAUDJ D., MCGURK G., MARGARSON S.,
 RA TEMPETE M., NORRIS V., HOLLAND I.B.;
 RL J. MOL. BIOL. 228:30-40(1992).
 DR EMBL; U02514; U02514.
 DR EMBL; X67470; ECGAMS.
 DR ECGENE; EG11545; FLGL.
 KW FLAGELLA.
 SQ SEQUENCE 317 AA; 34281 MW; 489131 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                || |||
LDKTNRGLKNSLNNVLT VRAELGT QLNELES LDSLGS DRALGQT QMSDLVD V DWNATISSYIM QQTALQAS
  230      240      250      260      270     280     290     300

```

```

YKAFTDM QGLSLF QLSK
      310

```

6. US-08-249-182-6 (1-9)

VAL1_PASV AL1 PROTEIN (ORF C1).

ID VAL1_PASV STANDARD; PRT; 323 AA.
 AC 000338;

RA 91157004
 RA DIZHOOR A.M., RAY S., KUMAR S., NIEMI G., SPENCER M., BROLLEY D.,
 RA WALSH K.A., PHILIPOV P.P., HURLEY J.B., STRYER L.;
 RL SCIENCE 251:915-918(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RM 92070564
 RA KUTUZOV M.A., SHMUKLER B.E., SUSLOV O.N., DERGACHEV A.E.,
 RA ZARGAROV A.A., ABDULAEV N.G.;
 RL FEBS LETT. 293:21-24(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RM 92335166
 RA RAY S., ZOZULYA S., NIEMI G.A., FLAHERTY K.M., BROLLEY D.,
 RA DIZHOOR A.M., MCKAY D.B., HURLEY J., STRYER L.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:5705-5709(1992).
 RN [4]
 RP SEQUENCE OF 17-35.
 RC TISSUE=RETINA;
 RM 91184124
 RA LAMBRECHT H.G., KOCH K.W.;
 RL EMBO J. 10:793-798(1991).
 RN [5]
 RP MYRISTOYLATION.
 RM 92355549
 RA DIZHOOR A.M., ERICSSON L.H., JOHNSON R.S., KUMAR S., OLSHEVSKAYA E.,
 RA ZOZULYA S., NEUBERT T.A., STRYER L., HURLEY J.B., WALSH K.A.;
 RL J. BIOL. CHEM. 267:16033-16036(1992).
 RN [6]
 RP RETRACTATION ON FUNCTION.
 RM 93248555
 RA HURLEY J.B., DIZHOOR A.M., STRYER L.;
 RL SCIENCE 260:740-740(1993).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY.
 RM 94061988
 RA FLAHERTY K.M., ZOZULYA S., STRYER L., MCKAY D.B.;
 RL CELL 75:709-716(1993).
 CC -!- FUNCTION: SEEMS TO BE IMPLICATED IN THE PATHWAY FROM RETINAL ROD
 CC GUANYLATE CYCLASE TO RHODOPSIN. MAY BE INVOLVED IN THE BLOCKING OF
 CC THE PHOSPHORYLATION OF RHODOPSIN.
 CC -!- BINDS TWO CALCIUM IONS; ONE WITH HIGH AFFINITY, THE OTHER WITH
 CC LOW AFFINITY.
 CC -!- PTM: THE N-TERMINAL GLYCINE IS LINKED TO ONE OF FOUR DIFFERENT
 CC TYPES OF ACYL GROUPS. THE MOST ABUNDANT IS MYRISTOLEATE (14:1),
 CC BUT 14:0, 14:2, AND 12:0 ACYL RESIDUES ARE ALSO PRESENT.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND
 CC MORE SPECIFICALLY TO THE RECOVERIN SUBFAMILY.
 DR EMBL; X63322; BTP26CBP.
 DR EMBL; S39582; S39582.
 DR PIR; A38433; A38433.
 DR PIR; S19305; S19305.
 DR PIR; A46129; A46129.
 DR PROSITE; PS00018; EF_HAND.
 KW CALCIUM-BINDING; MYRISTYLATION; VISION.
 FT INIT_MET 0 0
 FT LIPID 1 1 MYRISTATE.
 FT DOMAIN 36 47 ANCESTRAL CALCIUM SITE 1.
 FT CA_BIND 73 84 LOW AFFINITY.
 FT CA_BIND 109 120 HIGH AFFINITY.
 FT DOMAIN 159 170 ANCESTRAL CALCIUM SITE 4.
 FT CONFLICT 18 18 L -> Q (IN REF. 4).
 FT CONFLICT 20 20 T -> N (IN REF. 4).
 SQ SEQUENCE 201 AA; 23202 MW; 208090 CN;

3. US-08-249-182-6 (1-9)

RECO_MOUSE RECOVERIN (CANCER ASSOCIATED RETINOPATHY PROTEIN)

ID RECO_MOUSE STANDARD: PRT: 201 AA.
 AC P34057;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE RECOVERIN (CANCER ASSOCIATED RETINOPATHY PROTEIN) (CAR PROTEIN)
 DE (23 KD PHOTORECEPTOR CELL-SPECIFIC PROTEIN).
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RM 92339508
 RA MCGINNIS J.F., STEPANIK P.L., BAEHR W., SUBBARAYA I., LERIOUS V.;
 RL FEBS LETT. 302:172-176(1992).
 CC -!- FUNCTION: SEEMS TO BE IMPLICATED IN THE PATHWAY FROM RETINAL ROD
 CC GUANYLATE CYCLASE TO RHODOPSIN. MAY BE INVOLVED IN THE BLOCKING OF
 CC THE PHOSPHORYLATION OF RHODOPSIN.
 CC -!- BINDS TWO CALCIUM IONS; ONE WITH HIGH AFFINITY, THE OTHER WITH
 CC LOW AFFINITY.
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND
 CC MORE SPECIFICALLY TO THE RECOVERIN SUBFAMILY.
 DR EMBL; X66196; MM23KDA.
 KW CALCIUM-BINDING; MYRISTYLATION; VISION.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT DOMAIN 36 47 ANCESTRAL CALCIUM SITE 1 (BY SIMILARITY).
 FT CA_BIND 73 84 LOW AFFINITY (BY SIMILARITY).
 FT CA_BIND 109 120 HIGH AFFINITY (BY SIMILARITY).
 FT DOMAIN 159 170 ANCESTRAL CALCIUM SITE 4 (BY SIMILARITY).
 SQ SEQUENCE 201 AA; 23275 MW; 204261 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
DSDPKAYAQHVFERSFDANS DGTLD FKEYVIALHMTAGKPTQKLEWAFSLYDVDGNGTISKNEVLEIVMAIF
  60      70      80      90      100      110      X 120

KMIKPEDVKLLPDDENTPEKRAEKIWAFFGKKEDDKL
 130      140      150      160

```

4. US-08-249-182-6 (1-9)

RECO_BOVIN RECOVERIN (P26).

ID RECO_BOVIN STANDARD: PRT: 201 AA.
 AC P21457;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE RECOVERIN (P26).
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ARTIODACTYLA.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 5-194.
 RC TISSUE=RETINA;

2. US-08-249-182-6 (1-9)
RECO_HUMAN RECOVERIN (CANCER ASSOCIATED RETINOPATHY PROTEIN)

ID RECO_HUMAN STANDARD; PRT; 199 AA.
AC P35243;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE RECOVERIN (CANCER ASSOCIATED RETINOPATHY PROTEIN) (CAR PROTEIN).
GN RCV1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RM 92392330
RA MURAKAMI A., YAJIMA T., INANA G.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 187:234-244(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RM 93272873
RA WIECHMANN A.F., HAMMARBACK J.A.;
RL EXP. EYE RES. 56:463-470(1993).
RN [3]
RP SEQUENCE FROM N.A.
RM 92406381
RA THIRKILL C.E., TAIT R.C., TYLER N.K., ROTH A.M., KELTNER J.L.;
RL INVEST. OPHTHALMOL. VIS. SCI. 33:2768-2772(1992).
CC -!- FUNCTION: SEEMS TO BE IMPLICATED IN THE PATHWAY FROM RETINAL ROD
CC GUANYLATE CYCLASE TO RHODOPSIN. MAY BE INVOLVED IN THE BLOCKING OF
CC THE PHOSPHORYLATION OF RHODOPSIN.
CC -!- BINDS TWO CALCIUM IONS; ONE WITH HIGH AFFINITY, THE OTHER WITH
CC LOW AFFINITY.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND
CC MORE SPECIFICALLY TO THE RECOVERIN SUBFAMILY.
DR EMBL; S43855; S43855.
DR EMBL; S45545; S45545.
DR EMBL; S62028; S62028.
DR MIM; 179618; TENTH EDITION.
KW CALCIUM-BINDING; MYRISTYLATION; VISION.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 36 47 ANCESTRAL CALCIUM SITE 1 (BY SIMILARITY).
FT CA_BIND 73 84 LOW AFFINITY (BY SIMILARITY).
FT CA_BIND 109 120 HIGH AFFINITY (BY SIMILARITY).
FT DOMAIN 159 170 ANCESTRAL CALCIUM SITE 4 (BY SIMILARITY).
SQ SEQUENCE 199 AA; 22999 MW; 201221 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.69
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
||| | ||
DTPDKAYAGHVFRSFDNSLNDGTLDFKEYVIALHMTTAGKTNGKLEWAFSLYDVDGNGTISKNEVLEIVMAIF
60 70 80 90 100 110 X 120

KMITPEDVKLLPDDENTPEKRAEKIWKYFGKNDKDL
130 140 150 160

21. INB_HURSE	INTERFERON BETA PRECURSOR.	186	5	5	3.52	0
22. INB_HUMAN	INTERFERON BETA PRECURSOR (F1	187	5	5	3.52	0
23. CUP8_DRONE	PUPAL CUTICLE PROTEIN EDG-84	188	5	5	3.52	0
24. CHS2_AJECA	CHITIN SYNTHASE 2 (EC 2.4.1.1	194	5	5	3.52	0
25. YLH3_CAEEL	HYPOTHETICAL 25.6 KD PROTEIN	216	5	5	3.52	0
26. UCR2_TOBAC	CYTOCHROME B6-F COMPLEX IRDN-	228	5	5	3.52	0
27. UCR1_TOBAC	CYTOCHROME B6-F COMPLEX IRDN-	228	5	5	3.52	0
28. UCRI_PEA	CYTOCHROME B6-F COMPLEX IRDN-	230	5	5	3.52	0
29. SUMT_METIV	UROPORPHYRIN-III C-METHYLTRAN	230	5	6	3.52	0
30. GNTR_BACSU	GLUCONATE OPERON TRANSCRIPTIO	243	5	5	3.52	0
31. UCRI_SPIOL	CYTOCHROME B6-F COMPLEX IRDN-	247	5	5	3.52	0
32. HEMA_IAX3I	HEMAGGLUTININ PRECURSOR (FRAG	249	5	5	3.52	0
33. PMGY_ECOLI	PHOSPHOGLYCERATE MUTASE (EC 5	250	5	5	3.52	0
34. CLCA_PSEPU	CHLOROCLATECHOL 1,2-DIOXYGENAS	260	5	5	3.52	0
35. YGFD_ECOLI	HYPOTHETICAL 29.4 KD PROTEIN	268	5	5	3.52	0
36. TRPA_VIBPA	TRYPTOPHAN SYNTHASE ALPHA CHA	268	5	5	3.52	0
37. VG15_BPT4	TAIL CONNECTOR PROTEIN GP15.	272	5	5	3.52	0
38. TRA9_MYCTU	PUTATIVE TRANSPOSASE (INSERTI	278	5	5	3.52	0
39. DH47_ARATH	DEHYDRIN COR47 (COLD-INDUCED	294	5	5	3.52	0
40. XYL1_PICST	NAD(P)H-DEPENDENT XYLOSE REDU	318	5	5	3.52	0

1. US-08-249-182-6 (1-9)

VG01_BPP22 PORTAL PROTEIN (PROTEIN GP1).

ID VG01_BPP22 STANDARD; PRT; 724 AA.
AC P26744;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE PORTAL PROTEIN (PROTEIN GP1).
GN 1.
OS BACTERIOPHAGE P22.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.
RM 91306435
RA EPPLER K., WYCKOFF E., GOATES J., PARR R., CASJENS S.;
RL VIROLOGY 183:519-538(1991).
CC -!- FUNCTION: REQUIRED FOR SUCCESSFUL CONDENSATION OF DNA WITHIN THE
CC CAPSID. GP1 IS A MINOR STRUCTURAL PROTEIN. THE PORTAL PROTEIN IS
CC PRESENT AS A SINGLE RING-SHAPED DODECAMER LOCATED AT THE POINT
CC WHERE TAILS ATTACH. IT IS THROUGH THIS RING THAT DNA IS THOUGHT
CC TO ENTER THE PROHEAD.
CC -!- SUBUNIT: HOMODODECAMER.
DR EMBL; M59749; POP22PAC.
DR PIR; C40474; Z1BP22.
KW METAL-BINDING; LATE PROTEIN.
FT INIT_MET 0 0
FT METAL 149 149 POTENTIAL.
FT METAL 154 154 POTENTIAL.
FT METAL 171 171 POTENTIAL.
FT METAL 176 176 POTENTIAL.
SQ SEQUENCE 724 AA; 82611 MW; 2488846 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.86
Residue Identity = 77% Matches = 7 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
|||||||

TEAVNGGQVAFDTVNQLNMRADLETYYVFQDNLATAMRRDGEIYQSIIVNDIYDVPNRVTTITLEDGSEKDVQLM
430 440 450 460 470 480 490

AEVVDLATGEKQVLNDIRGRYECYTDVGPSFQSMKQ

SCORE 0 1 2 3 4 5 6 7
STDEV -1 0 1 2 4 5

PARAMETERS

Similarity matrix Unitary K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 5
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

Initial scores to save 40 Alignments to save 15
Optimized scores to save 0 Display context 50

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
 2 3 0.85

Times: CPU Total Elapsed
 00:00:50.95 00:01:02.00

Number of residues: 12496420
Number of sequences searched: 36000
Number of scores above cutoff: 4523

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

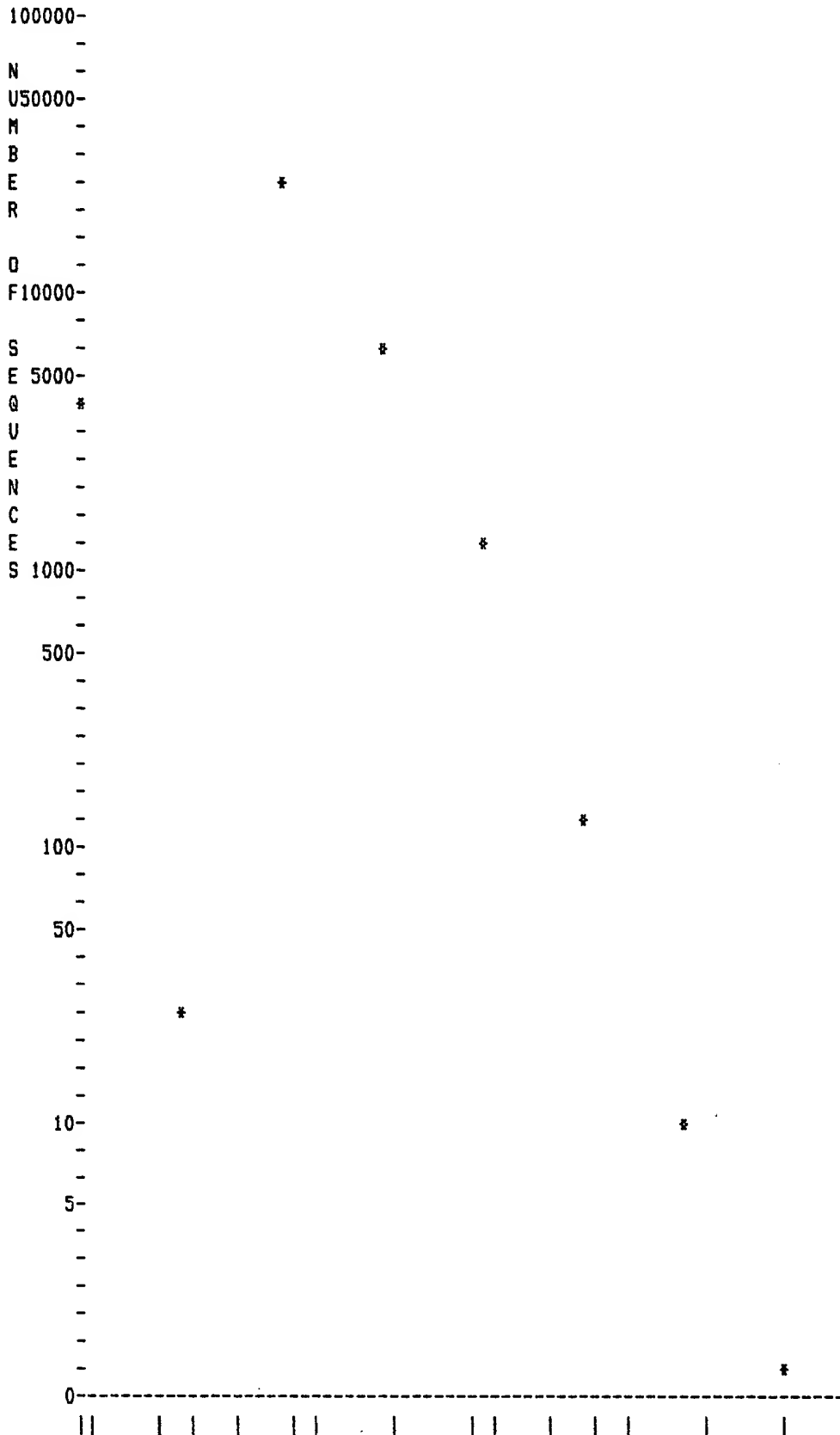
Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
1. VG01_BPP22	PORTAL PROTEIN (PROTEIN GP1).	724	7	7	5.86	0
**** 4 standard deviations above mean ****						
2. RECO_HUMAN	RECOVERIN (CANCER ASSOCIATED	199	6	6	4.69	0
3. RECO_MOUSE	RECOVERIN (CANCER ASSOCIATED	201	6	6	4.69	0
4. RECO_BOVIN	RECOVERIN (P26).	201	6	6	4.69	0
5. FLGL_ECOLI	FLAGELLAR HOOK-ASSOCIATED PRO	317	6	6	4.69	0
6. VAL1_PASV	AL1 PROTEIN (ORF C1).	323	6	6	4.69	0
7. ATN2_YEAST	SODIUM TRANSPORT ATPASE 2 (EC	1091	6	6	4.69	0
8. ATN1_YEAST	SODIUM TRANSPORT ATPASE 1 (EC	1091	6	6	4.69	0
9. CYGS_STRPU	SPERACT RECEPTOR PRECURSOR (G	1125	6	6	4.69	0
10. YDG7_CAEEL	HYPOTHETICAL 375.7 KD PROTEIN	3343	6	6	4.69	0
**** 3 standard deviations above mean ****						
11. TES1_RAT	TESTIN 1 (CMB-22) (FRAGMENT).	30	5	5	3.52	0
12. VKIL_BPP22	KIL PROTEIN.	62	5	5	3.52	0
13. FER1_EQUITE	FERREDOXIN I.	95	5	5	3.52	0
14. FER1_EQUAR	FERREDOXIN I.	95	5	5	3.52	0
15. CYTC_BOVIN	CYSTATIN, COLOSTRUM (THIOL PR	112	5	6	3.52	0
16. RS24_STRPU	40S RIBOSOMAL PROTEIN S24.	130	5	5	3.52	0
17. CYT1_MAIZE	CYSTATIN I PRECURSOR (CORN KE	135	5	5	3.52	0
18. LEPA_PSEFL	LEPA PROTEIN (FRAGMENT).	161	5	5	3.52	0
19. VGG_BPG4	MAJOR SPIKE PROTEIN (G PROTEI	177	5	5	3.52	0
20. INB_MOUSE	INTERFERON BETA PRECURSOR.	182	5	5	3.52	0

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_6s.res made by on Thu 22 Sep 94 10:11:13-PDT.

Query sequence being compared:US-08-249-182-6 (1-9)
Number of sequences searched: 36000
Number of scores above cutoff: 4523

Results of the initial comparison of US-08-249-182-6 (1-9) with:
Data bank : Swiss-Prot 28, all entries



ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change
 30-Sep-1993
 ACCESSIONS A34199; PC1250
 REFERENCE A92740
 #authors Cheng, C.Y.; Grima, J.; Stahler, M.S.; Lockshin, R.A.;
 Bardin, C.W.
 #journal J. Biol. Chem. (1989) 264:21386-21393
 #title Testins are structurally related sertoli cell proteins whose
 secretion is tightly coupled to the presence of germ cells.
 #cross-references MUID:90078247
 #accession A34199
 ##molecule_type protein
 ##residues 1-30 ##label CHE
 SUMMARY #length 30 #checksum 6050
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.22
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
      || |||
TPDPSLDVEWNEWRTKHGKTYNMNEERLKR
      X 10  X   20   30
  
```

15. US-08-249-182-6 (1-9)

JU0054 hypothetical adh1 protein - Clostridium acetobutylicum

ENTRY JU0054 #type fragment
 TITLE hypothetical adh1 protein - Clostridium acetobutylicum
 (fragment)
 ORGANISM #formal_name Clostridium acetobutylicum
 DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
 30-Sep-1993
 ACCESSIONS JU0054
 REFERENCE A91610
 #authors Youngleson, J.S.; Jones, W.A.; Jones, D.T.; Woods, D.R.
 #journal Gene (1989) 78:355-364
 #title Molecular analysis and nucleotide sequence of the adh1 gene
 encoding an NADPH-dependent butanol dehydrogenase in the
 gram-positive anaerobe Clostridium acetobutylicum.
 #cross-references MUID:89378762
 #accession JU0054
 ##molecule_type DNA
 ##residues 1-52 ##label YDU
 ##note the amino acid sequence is not shown in this paper
 GENETICS
 #gene adh1
 SUMMARY #length 52 #checksum 6532
 SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.22
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
      || |||
DLIGLDVCLAIMDVLFNETGDSKYRASSILRKYVRAGWLGKSGKGFYDYSK
      10 X   20   30   40   50
  
```

KEYWORDS

cyclase catalytic domain homology; protein kinase homology
carbon-oxygen lyase; cGMP synthesis; glycoprotein; hormone
receptor; membrane protein; phosphorus-oxygen lyase

FEATURE

1-21 #domain signal sequence #status predicted #label SIG\
22-1125 #protein speract receptor #status predicted #label MAT\
22-510 #domain extracellular #status predicted #label EXT\
509-530 #domain transmembrane #status predicted #label TMM\
532-1125 #domain intracellular #status predicted #label CYT\
573-841 #domain protein kinase homology #label KIN\
892-1093 #domain guanylate cyclase catalytic domain homology
#label CAT

SUMMARY

#length 1125 #molecular-weight 126256 #checksum 5278

SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| |||
IAVLPETFEMVSIFFSIDIVGFTALSAASTPIQVVNLLNDLYTLFDAIISNYDVYKVETIGDAYMLVSGCLPLR
  910      920      930      940      950 X      960      970

NGDRHAGQIASTAHLLSVKGFIVPHKPEVFLKLRI
  980      990     1000     1010

```

13. US-08-249-182-6 (1-9)

PC1251 testin II - rat (fragment)

ENTRY PC1251 #type fragment
TITLE testin II - rat (fragment)
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
30-Sep-1993

ACCESSIONS

PC1251

REFERENCE

PC1250

#authors Cheng, C.Y.; Morris, I.; Bardin, C.W.
#journal Biochem. Biophys. Res. Commun. (1993) 191:224-231
#title Testins are structurally related to the mouse cysteine
proteinase precursor but devoid of any
protease/anti-protease activity.

#accession PC1251

##molecule_type protein

##residues 1-19 ##label CHE

SUMMARY

#length 19 #checksum 4922

SEQUENCE

Initial Score = 5 Optimized Score = 5 Significance = 4.22
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

      X      X
      YDVPWNETI
      || |||
AAPTDPDPSLDVEWNEWRTK
  10      X

```

14. US-08-249-182-6 (1-9)

A34199 testin I - rat (fragment)

ENTRY A34199 #type fragment
TITLE testin I - rat (fragment)

TITLE Ca2+-transporting ATPase (EC 3.6.1.38) PMR2 - yeast
 (Saccharomyces cerevisiae)
 ALTERNATE_NAMES calcium pump PMR2
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change
 05-May-1994
 ACCESSIONS S05788; B30990
 REFERENCE A30990
 #authors Rudolph, H.K.; Antebi, A.; Fink, G.R.; Buckley, C.M.; Dornan,
 T.E.; LeVitre, J.; Davidow, L.S.; Mao, J.I.; Moir, D.T.
 #journal Cell (1989) 58:133-145
 #title The yeast secretory pathway is perturbed by mutations in
 PMR1, a member of a Ca(2+) ATPase family.
 #cross-references MUID:89324047
 #accession S05788
 ##molecule_type DNA
 ##residues 1-1091 ##label RUD
 ##cross-references ENBL:M25489
 GENETICS
 #gene LISTA:PMR2
 #map_position 4R
 CLASSIFICATION #superfamily Na+/K+-transporting ATPase alpha chain
 KEYWORDS ATP; calcium transport; hydrolase; membrane protein
 FEATURE
 369 #active_site Asp #status predicted
 SUMMARY #length 1091 #molecular-weight 120356 #checksum 4741
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                || |||
YSLSNEGLRVLGFASKSFTKDQVNDQ LKNITSNRATAESDLVFLGLIGIYDPPRNETAGAVKKFHQAGINV
600      610      620      630      640      650      X 660

HMLTGDFVGTAKAIAQEVGILPTNLYHYSQEI VDSMV
670      680      690      700

```

12. US-08-249-182-6 (1-9)

DYURCP speract receptor precursor - sea urchin (Strongylo

ENTRY DYURCP #type complete
 TITLE speract receptor precursor - sea urchin (Strongylocentrotus
 purpuratus)
 ALTERNATE_NAMES guanylate cyclase, membrane-bound
 CONTAINS guanylate cyclase (EC 4.6.1.2)
 ORGANISM #formal_name Strongylocentrotus purpuratus #common_name
 purple urchin
 DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
 31-Dec-1993
 ACCESSIONS A33535; A30856
 REFERENCE A33535
 #authors Thorpe, D.S.; Garbers, D.L.
 #journal J. Biol. Chem. (1989) 264:6545-6549
 #title The membrane form of guanylate cyclase. Homology with a
 subunit of the cytoplasmic form of the enzyme.
 #cross-references MUID:89197965
 #accession A33535
 ##molecule_type mRNA
 ##residues 1-1125 ##label THD
 ##cross-references GB:J04693
 CLASSIFICATION #superfamily membrane-bound guanylate cyclase; guanylate

Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| ||
TPGIVAGYPARMMIKLLYQLTEPEEPSHYISMLPKLAKVPNVVQEHVKPRSDVPWSETVEELDWREALAILR
680      690      700      710      720      730      740      750

HRPARTSELVTQWLQLIGRFTAHPDKRALTLRYWF
      760      770      780

```

10. US-08-249-182-6 (1-9)

S25007 Ca2+-transporting ATPase (EC 3.6.1.38) ENA2 - yeas

```

ENTRY      S25007      #type complete
TITLE      Ca2+-transporting ATPase (EC 3.6.1.38) ENA2 - yeast
            (Saccharomyces cerevisiae)
ORGANISM    #formal_name Saccharomyces cerevisiae
DATE        07-May-1993 #sequence_revision 07-May-1993 #text_change
            28-May-1993
ACCESSIONS  S25007; S30898
REFERENCE   S25007
            #authors    Rodriguez-Navarro, A.
            #submission  submitted to the EMBL Data Library, June 1992
            #accession   S25007
            ##molecule_type DNA
            ##residues    1-1091 ##label ROD
            ##cross-references EMBL:X67136
REFERENCE   S30898
            #authors    Garciadeblas, B.; Rubio, F.; Quintero, F.J.; Banuelos, M.A.;
            Haro, R.; Rodriguez-Navarro, A.
            #journal     Mol. Gen. Genet. (1993) 236:363-368
            #title       Differential expression of two genes encoding isoforms of the
            ATPase involved in sodium efflux in Saccharomyces
            cerevisiae.
            #accession   S30898
            ##molecule_type DNA
            ##residues    1-97;752-754;1088-1091 ##label R02
            ##cross-references EMBL:X67136
GENETICS
            #gene        ENA2
            #map_position 4
CLASSIFICATION #superfamily Na+/K+-transporting ATPase alpha chain
KEYWORDS       calcium transport; hydrolase; ion transport; membrane protein
SUMMARY        #length 1091 #molecular-weight 120316 #checksum 5270
SEQUENCE

```

Initial Score = 6 Optimized Score = 6 Significance = 5.28
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                || | ||
YLSNEGLRVLGFAKSFYTKDQVNDQDLKNITSNRATAESDLVFLGLIGIYDPPRNETAGAVKKFHQAGINV
600      610      620      630      640      650      X 660

HMLTGDFVGTAKAIAQEVGILPTNLYHYSQEIYDSMV
670      680      690      700

```

11. US-08-249-182-6 (1-9)

PWBYR2 Ca2+-transporting ATPase (EC 3.6.1.38) PHR2 - yeas

```

ENTRY      PWBYR2      #type complete

```

ENTRY J01552 #type complete
 TITLE C1 protein - Panicum streak virus
 ORGANISM #formal_name Panicum streak virus
 DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
 30-Sep-1993
 ACCESSIONS J01552
 REFERENCE J01549
 #authors Briddon, R.W.; Lunness, P.; Chamberlin, L.C.L.; Pinner, M.S.;
 Brundish, H.; Markham, P.G.
 #journal J. Gen. Virol. (1992) 73:1041-1047
 #title The nucleotide sequence of an infections insect-transmissible
 clone of the geminivirus Panicum streak virus.
 #contents Isolate Kenya
 #accession J01552
 ##molecule_type DNA
 ##residues 1-323 ##label BRI
 ##cross-references EMBL:X60168
 CLASSIFICATION #superfamily tomato golden mosaic virus AL1 protein
 SUMMARY #length 323 #molecular-weight 37026 #checksum 1817
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||  |||
HATSREEYLSLVQSSLPYDWATKLNIFYEVSASRLFPDIAEPYTNPHPTTEYDLHCNETIEDWLKPNIYQVSP
  160      170      180      190      200 X      210      220

GAYKLLPSCLSLEQAIADLEWLDDTTRMLQEKEREA
  230      240      250      260

```

9. US-08-249-182-6 (1-9)

B47521 putative RNA-dependent RNA polymerase fusion prote

ENTRY B47521 #type complete
 TITLE putative RNA-dependent RNA polymerase fusion protein -
 giardiavirus GLV
 ORGANISM #formal_name giardiavirus, GLV
 DATE 21-Jan-1994; #sequence_revision 21-Jan-1994; #text_change
 21-Jan-1994
 ACCESSIONS B47521
 REFERENCE A47521
 #authors Wang, A.L.; Yang, H.M.; Shen, K.A.; Wang, C.C.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:8595-8599
 #title Giardiavirus double-stranded RNA genome encodes a capsid
 polypeptide and a gag-pol-like fusion protein by a
 translation frameshift.
 #cross-references MUID:93391401
 #contents host= Giardia lamblia
 #accession B47521
 ##status preliminary
 ##molecule_type nucleic acid
 ##residues 1-1057 ##label WAN
 ##cross-references NCBIN:137593; NCBIP:137595
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 1057 #molecular-weight 120098 #checksum 72
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
 Residue Identity = 66% Matches = 6 Mismatches = 3

Zargardov, A.A.; Abdulaev, N.G.

#journal FEBS Lett. (1991) 293:21-24
#title P26 - calcium binding protein from bovine retinal
photoreceptor cells.
#cross-references MUID:92070564
#accession S19305
##status preliminary
##residues 1-202 ##label KUT
##cross-references EMBL:X63322

SUMMARY #length 202 #molecular-weight 23333 #checksum 1961

SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
EADPKAYAGHVFRSFDANS DGTLD FKEYVIALHMTSAGKT NQKLEWAFSLYD VDGNGTISKNEVLEIVTAIF
60      70      80      90      100     110     120     130

KMISPEDTKHLPEDENTPEKRAEKIWGFFGKKDDDKL
      140      150      160
```

7. US-08-249-182-6 (1-9)

A38433 recoverin - bovine

ENTRY A38433 #type complete
TITLE recoverin - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change
30-Sep-1993
ACCESSIONS A38433
REFERENCE A38433
#authors Dizhoor, A.M.; Ray, S.; Kumar, S.; Niemi, G.; Spencer, M.;
Brolley, D.; Walsh, K.A.; Philipov, P.P.; Hurley, J.B.;
Stryer, L.
#journal Science (1991) 251:915-918
#title Recoverin: a calcium sensitive activator of retinal rod
guanylate cyclase.
#cross-references MUID:91157004
#accession A38433
##status preliminary
##molecule_type mRNA
##residues 1-202 ##label DIZ
##cross-references GB:M95858; GB:M77094

SUMMARY #length 202 #molecular-weight 23333 #checksum 1961

SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
EADPKAYAGHVFRSFDANS DGTLD FKEYVIALHMTSAGKT NQKLEWAFSLYD VDGNGTISKNEVLEIVTAIF
60      70      80      90      100     110     120     130

KMISPEDTKHLPEDENTPEKRAEKIWGFFGKKDDDKL
      140      150      160
```

8. US-08-249-182-6 (1-9)

Initial Score = 6 Optimized Score = 6 Significance = 5.28
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
DSDPKAYAQHVFERSFDANS DGTLD FKEYVIALHMTTAGKPTQKLEWAFSLYDVDGNGTISKNEVLEIVMAIF
60      70      80      90      100     110     120     130

KMIKPEDVKLLPDDENTPEKRAEKIWAFFGKKEDDKL
      140      150      160
```

5. US-08-249-182-6 (1-9)

A46129 recoverin=calcium sensor - bovine

ENTRY A46129 #type complete
TITLE recoverin=calcium sensor - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 21-Sep-1993; #sequence_revision 21-Sep-1993; #text_change 21-Sep-1993
ACCESSIONS A46129
REFERENCE A46129
#authors Ray, S.; Zozulya, S.; Niemi, G.A.; Flaherty, K.M.; Brolley, D.; Dizhoor, A.M.; McKay, D.B.; Hurley, J.; Stryer, L.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:5705-5709
#title Cloning, expression, and crystallization of recoverin, a calcium sensor in vision.
#cross-references MUID:92335166
#contents retinas
#accession A46129
##status preliminary
##molecule_type mRNA
##residues 1-202 ##label RAY
##cross-references NCBIN:108412; NCBIP:108413
##note sequence extracted from NCBI backbone
SUMMARY #length 202 #molecular-weight 23333 #checksum 1961
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
Residue Identity = 66% Matches = 6 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
EADPKAYAQHVFERSFDANS DGTLD FKEYVIALHMTSAGKTNQKLEWAFSLYDVDGNGTISKNEVLEIVTAIF
60      70      80      90      100     110     120     130

KMISPEDTKHLPEDENTPEKRAEKINGFFGKKDDDKL
      140      150      160
```

6. US-08-249-182-6 (1-9)

S19305 calcium-binding protein P26 - bovine

ENTRY S19305 #type complete
TITLE calcium-binding protein P26 - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993
ACCESSIONS S19305
REFERENCE S19305
#authors Kutuzov, M.A.; Shmukler, B.E.; Suslov, O.N.; Dergachev, A.E.;

3. US-08-249-182-6 (1-9)

JC1227 recoverin - human

ENTRY JC1227 #type complete
 TITLE recoverin - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
 04-Mar-1994
 ACCESSIONS JC1227
 REFERENCE JC1227
 #authors Murakami, A.; Yajima, T.; Inana, G.
 #journal Biochem. Biophys. Res. Commun. (1992) 187:234-244
 #title Isolation of human retinal genes: recoverin cDNA and gene.
 #cross-references MUID:92392330
 #accession JC1227
 ##molecule_type mRNA
 ##residues 1-200 ##label MUR
 COMMENT This protein modulates guanylate cyclase activity.
 GENETICS
 #gene GDB:RCV1
 #map_position 17
 #introns 127/3; 165/1
 FEATURE
 37-47 #domain calcium-binding (EF-hand) #label CB1\
 73-84 #domain calcium-binding (EF-hand) #label CB2\
 108-122 #domain calcium-binding (EF-hand) #label CB3
 SUMMARY #length 200 #molecular-weight 23130 #checksum 7845
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 5.28
 Residue Identity = 66% Matches = 6 Mismatches = 3
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                ||| | ||
DTDPKAYAQHVFERSFDSNLDGTLDFKEYVIALHMTTAGKTNGKLEWAFSLYDVGNGTISKNEVLEIVMAIF
60      70      80      90      100      110      120      130

KMITPEDVKLLPDDENTPEKRAEKINKYFGKNDDDKL
      140      150      160

```

4. US-08-249-182-6 (1-9)

S21155 23K protein - mouse

ENTRY S21155 #type complete
 TITLE 23K protein - mouse
 ORGANISM #formal_name Mus musculus #common_name house mouse
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
 22-Nov-1993
 ACCESSIONS S21155
 REFERENCE S21155
 #authors McGinnis, J.F.; Stepanik, P.L.; Baehr, W.; Subbaraya, I.;
 Lerious, V.
 #journal FEBS Lett. (1992) 302:172-176
 #title Cloning and sequencing of the 23 kDa mouse photoreceptor
 cell-specific protein.
 #cross-references MUID:92339508
 #accession S21155
 ##status preliminary
 ##residues 1-202 ##label MCG
 SUMMARY #length 202 #molecular-weight 23406 #checksum 526
 SEQUENCE

#title Identification, purification, and partial sequence analysis
 of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329
 ##status preliminary
 ##molecule_type protein
 ##residues 1-114 ##label STR
 ##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
 NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
 NCBIP:78509; NCBIP:78508; NCBIP:78503
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 114 #checksum 7335
 SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 8.44
 Residue Identity = 100% Matches = 9 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
RDIEHLTSLDFFRVNSMGTVFVGVGPTFKGGQPLWITATKSPPFENINLYYDVPWNETIPEEVTXPNYLQAE
   30      40      50      60      70 X      80      90

VSYPAFKXPXLDVYKWHVAAN
   100      110
  
```

2. US-08-249-182-6 (1-9)

Z1BP22 gene 1 protein - phage P22

ENTRY Z1BP22 #type complete
 TITLE gene 1 protein - phage P22
 ORGANISM #formal_name phage P22
 #note host Salmonella typhimurium
 DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
 08-Apr-1994
 ACCESSIONS C40474
 REFERENCE A40474
 #authors Eppler, K.; Wyckoff, E.; Goates, J.; Parr, R.; Casjens, S.
 #journal Virology (1991) 183:519-538
 #title Nucleotide sequence of the bacteriophage P22 genes required
 for DNA packaging.
 #cross-references MUID:91306435
 #accession C40474
 ##molecule_type DNA
 ##residues 1-725 ##label EPP
 ##cross-references GB:M59749

GENETICS

#gene 1
 CLASSIFICATION #superfamily phage P22 gene 1 protein
 SUMMARY #length 725 #molecular-weight 82742 #checksum 9976
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 6.33
 Residue Identity = 77% Matches = 7 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
TEAVNGGQVAFDVTNQLNMRADLETYVFQDNLATAMRRDGEIYQSI VNDIYDVP RNVTTITLEDGSEKDVQLM
   430      440      450      460      470 X      480      490

AEVVDLATGEKQVLNDIRGRYECYTDVGPFSQSMKQQ
   500      510      520      530
  
```

Sequence Name	Description	Length	Score	Score	Sig.	Frame
1. A42329	autotaxin - human (fragments)	114	9	9	8.44	0

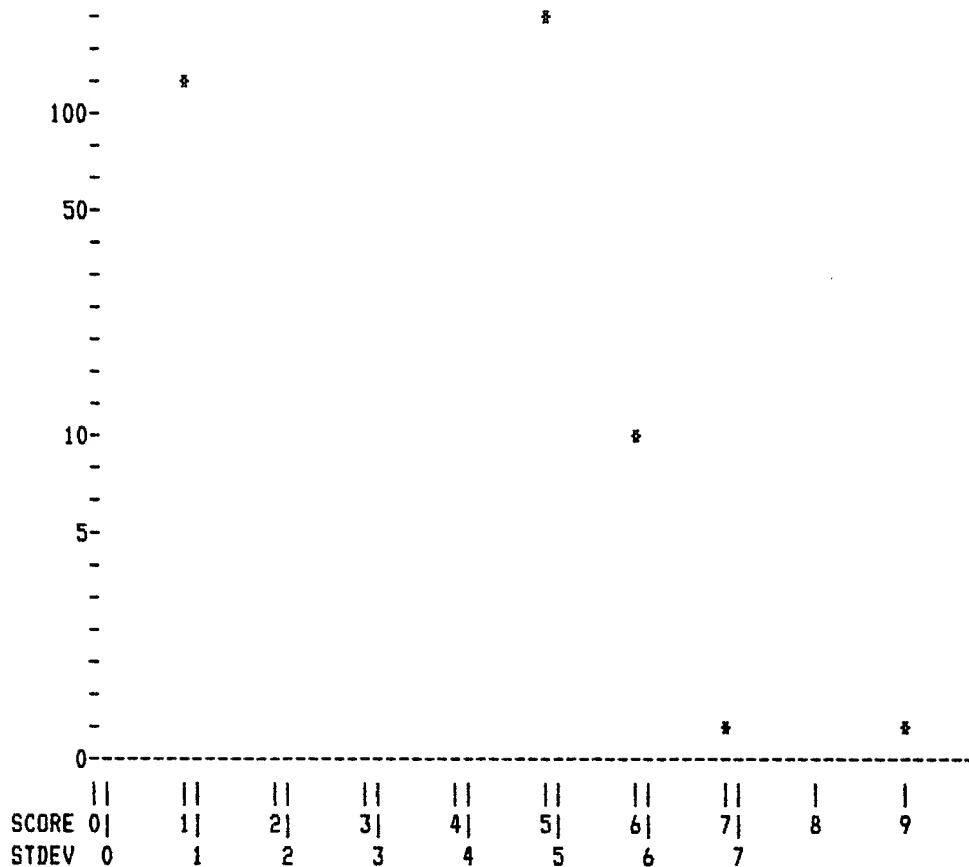
The list of other best scores is:

Sequence Name	Description	Length	Score	Score	Sig.	Frame
**** 6 standard deviations above mean ****						
2. Z1BP22	gene 1 protein - phage P22	725	7	7	6.33	0
**** 5 standard deviations above mean ****						
3. JC1227	recoverin - human	200	6	6	5.28	0
4. S21155	23K protein - mouse	202	6	6	5.28	0
5. A46129	recoverin=calcium sensor - bo	202	6	6	5.28	0
6. S19305	calcium-binding protein P26 -	202	6	6	5.28	0
7. A38433	recoverin - bovine	202	6	6	5.28	0
8. J01552	C1 protein - Panicum streak v	323	6	6	5.28	0
9. B47521	putative RNA-dependent RNA po	1057	6	6	5.28	0
10. S25007	Ca2+-transporting ATPase (EC	1091	6	6	5.28	0
11. FWBYR2	Ca2+-transporting ATPase (EC	1091	6	6	5.28	0
12. OYURCP	speract receptor precursor -	1125	6	6	5.28	0
**** 4 standard deviations above mean ****						
13. PC1251	testin II - rat (fragment)	19	5	5	4.22	0
14. A34199	testin I - rat (fragment)	30	5	5	4.22	0
15. JU0054	hypothetical adh1 protein - C	52	5	5	4.22	0
16. VDBP22	kil protein - phage P22	62	5	5	4.22	0
17. FEE01F	ferredoxin I - field horsetai	95	5	5	4.22	0
18. FEE01	ferredoxin I - horsetail (Equ	95	5	5	4.22	0
19. UDB0	cystatin - bovine	112	5	6	4.22	0
20. S36733	malate dehydrogenase (NADP+)	125	5	5	4.22	0
21. A43907	ribosomal protein S24 - sea u	130	5	5	4.22	0
22. S27239	cysteine proteinase inhibitor	135	5	5	4.22	0
23. A45708	immunoreactive epitope (overl	144	5	5	4.22	0
24. S36680	lepA protein - Pseudomonas fl	161	5	5	4.22	0
25. S04479	Interferon beta-1 - Human	166	5	5	4.22	0
26. A23537	gnt operon regulatory protein	169	5	5	4.22	0
27. ZGBPG4	gene G protein - phage G4	177	5	5	4.22	0
28. S02020	interferon beta precursor - m	182	5	5	4.22	0
29. IVMSB	interferon beta precursor - m	182	5	5	4.22	0
30. IVH0B1	interferon beta-I precursor -	186	5	5	4.22	0
31. IVHUB1	interferon beta-1 precursor -	187	5	5	4.22	0
32. B49773	ecdysone-dependent cuticle pr	188	5	5	4.22	0
33. H45189	chitin synthase=HcCHS2 gene p	195	5	5	4.22	0
34. S38667	ribonuclease S3 - Peruvian to	195	5	5	4.22	0
35. S34833	stylar protein - Peruvian tom	205	5	5	4.22	0
36. S40720	hypothetical protein - Caenor	216	5	5	4.22	0
37. S39545	plastoquinol--plastocyanin re	228	5	5	4.22	0
38. S25312	plastoquinol--plastocyanin re	228	5	5	4.22	0
39. S23735	plastoquinol--plastocyanin re	228	5	5	4.22	0
40. S26199	plastoquinol--plastocyanin re	230	5	5	4.22	0

1. US-08-249-182-6 (1-9)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
 Cioco, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	0.95

Times:	CPU	Total Elapsed
	00:01:25.89	00:01:50.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	3789

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

PI anti-neoplastic and immuno-modulatory agents.

PS Disclosure; Fig 14; 96pp; English.

CC The sequence encodes IFN-beta(IFN-beta(2-7)-IFN-alpha2(1-5)-

CC (IFN-beta(9-56)-IFN-alpha1(7-54)) a modified interferon-beta which has

CC increased biological activity compared to natural IFN-beta, and

CC which is more effective in the treatment of viral or neoplastic

CC diseases, or immunosuppressed or immunodeficient conditions.

SQ Sequence 165 AA;

SQ 7 A; 12 R; 10 N; 7 D; 0 B; 3 C; 8 Q; 11 E; 0 Z; 6 G; 7 H;

SQ 11 I; 21 L; 8 K; 6 M; 10 F; 4 P; 11 S; 8 T; 2 W; 8 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34

Residue Identity = 55% Matches = 5 Mismatches = 4

Gaps = 0 Conservative Substitutions = 0

X X

YDVPWNETI

||||

RISPSSCLMDRHDGFGPQEEFDGNQGFQKAPAILTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN

30 40 50 60 70 X 80 X 90

HLKTVLEEKLEKEDFTRGKLMSSLHLKRYYGRIHLHYL

100 110 120 130

> D <

0| |0 IntelliGenetics

> D <

FastDB - Fast Pairwise Comparison of Sequences

Release 5.4

Results file u249_6p.res made by on Thu 22 Sep 94 10:38:59-PDT.

Query sequence being compared:US-08-249-182-6 (1-9)

Number of sequences searched: 70848

Number of scores above cutoff: 3789

Results of the initial comparison of US-08-249-182-6 (1-9) with:

Data bank : PIR 41, all entries

100000-

-

N -

U50000-

*

M -

B -

E -

R -

-

D *

*

F10000-

-

S -

E 5000-

Q -

U -

E -

N -

*

C -

E -

S 1000-

-

-

500-

-

-

HLKTCLEEKLEKEDFTRGKLMSSLHLKRYYGRIILHYL

100 110 120 130

14. US-08-249-182-6 (1-9)

P50280 Protein sequence encoding modified interferon-beta

ID P50280 standard; protein; 165 AA.
AC P50280;
DT 09-DEC-1991 (first entry)
DE Protein sequence encoding modified interferon-beta.
KW Interferon-beta; IFNX403; virucide; antitumor; immunostimulant.
OS Synthetic.
PN EP-131816-A.
PD 23-JAN-1985.
PF 28-JUN-1984; 107498.
PR 01-JUL-1983; GB-017880.
PA (SEAR) SEARLE G D & CO.
PI Bell LD; Boseley PG; Smith JC; Houghton M.
DR WPI; 85-020165/04.
DR N-PSDB; N50307.
PT New modified beta-interferon(s) - useful as antiviral,
PT anti-neoplastic and immuno-modulatory agents.
PS Disclosure; Fig 22; 96pp; English.
CC The sequence encodes IFN-beta(IFN-beta(9-56)-IFN-alpha2(7-53)
CC (Leu17-Cys17) a modified interferon-beta which has
CC increased biological activity compared to natural IFN-beta, and
CC which is more effective in the treatment of viral or neoplastic
CC diseases, or immunosuppressed or immunodeficient conditions.
SQ Sequence 165 AA;
SQ 6 A; 12 R; 10 N; 4 D; 0 B; 3 C; 7 Q; 12 E; 0 Z; 8 G; 7 H;
SQ 11 I; 22 L; 10 K; 5 M; 11 F; 1 P; 11 S; 9 T; 2 W; 9 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
|||||
RKISLFSCLKDRHDFGFPQEEFGNQFQKAETILTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN
30 40 50 60 70 X 80 X 90

HLKTVLEEKLEKEDFTRGKLMSSLHLKRYYGRIILHYL
100 110 120 130

15. US-08-249-182-6 (1-9)

P50273 Protein sequence encoding modified interferon-beta

ID P50273 standard; protein; 165 AA.
AC P50273;
DT 09-DEC-1991 (first entry)
DE Protein sequence encoding modified interferon-beta.
KW Interferon-beta; IFNX410; virucide; antitumor; immunostimulant.
OS Synthetic.
PN EP-131816-A.
PD 23-JAN-1985.
PF 28-JUN-1984; 107498.
PR 01-JUL-1983; GB-017880.
PA (SEAR) SEARLE G D & CO.
PI Bell LD; Boseley PG; Smith JC; Houghton M.
DR WPI; 85-020165/04.
DR N-PSDB; N50300.
PT New modified beta-interferon(s) - useful as antiviral,

CC N301577. HinfI was used to digest the DNA sequences in the region
 CC of significant handicaps (see N30153, N30154, N30158, N30159), and
 CC the restriction fragments were ligated to form hybrid DNA.
 SQ Sequence 165 AA;
 SQ 7 A; 12 R; 9 N; 6 D; 0 B; 3 C; 10 Q; 11 E; 0 Z; 6 G; 7 H;
 SQ 12 I; 20 L; 10 K; 6 M; 9 F; 3 P; 10 S; 9 T; 2 W; 8 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
RISPFSCCLKDRHDFGFPQEIQLQGFQKEDAALTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN
  30      40      50      60      70 X   80 X   90

HLKTVLEEKLEKEDFTRGKLMSSSLHKRYYYGRILHYL
  100     110     120     130

```

13. US-08-249-182-6 (1-9)

P50207 Modified human interferon (IFN-457).

ID P50207 standard; protein; 165 AA.
 AC P50207;
 DT 01-JAN-1980 (first entry)
 DE Modified human interferon (IFN-457).
 KW Interferon; antitumor; virucide; immunostimulant;
 KW protein engineering; ss.
 OS Homo sapiens.
 PN EP-146413-A.
 PD 26-JUN-1985.
 PF 20-DEC-1984; 308984.
 PR 21-DEC-1983; GB-034102.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Smith JC, Porter AG, Adair JR;
 DR WPI; 85-154456/26.
 DR N-PSDB; N50214.
 PT New interferon with modified cysteine pattern - useful for
 PT altered antiviral and anti-proliferative effects, stability etc.
 PS Disclosure; Page 54; 62pp; English.
 CC The IFN is obtained by expression from a microorganism which has been
 CC transformed by recombinant DNA techniques. The modified IFN may
 CC allow selective properties to be shown, e.g. in the alteration of
 CC cellular membranes, as virucidal, immunomodulatory, antitumor and an
 CC antiproliferative agent. It has increased stability which results
 CC in improved recovery during production, increased storage life and
 CC prolonged activity. The IFN may be human-alpha + human-beta, human-
 CC beta + human-alpha or human-beta + human-gamma. Cysteine deletions
 CC or substitutions permit the formation of disulfide linkages at
 CC positions corresponding to linkages in a 2nd IFN. See also
 CC N50207-13 and P50192-97 and P50206.
 SQ Sequence 165 AA;
 SQ 6 A; 11 R; 11 N; 6 D; 0 B; 4 C; 12 Q; 13 E; 0 Z; 5 G; 5 H;
 SQ 11 I; 23 L; 11 K; 4 M; 9 F; 2 P; 9 S; 7 T; 3 W; 9 Y; 4 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
NGRLEYCLKDRMNFDPPEEIQLQGFQKEDAALTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN
  30      40      50      60      70 X   80 X   90

```

PA 01-JUL-1983; GB-017880.
 PI (SEAR) SEARLE G D & CO.
 PI Bell LD; Boseley PG; Smith JC; Houghton M.
 DR WPI; 85-020165/04.
 DR N-PSDB; N50308.
 PT New modified beta-interferon(s) - useful as antiviral,
 PT anti-neoplastic and immuno-modulatory agents.
 PS Disclosure; Fig 21; 96pp; English.
 CC The sequence encodes IFN-beta(IFN-beta(1-56)-IFN-alpha2(1-53)
 CC (Leu16-Cys16) a modified interferon-beta which has
 CC increased biological activity compared to natural IFN-beta, and
 CC which is more effective in the treatment of viral or neoplastic
 CC diseases, or immunosuppressed or immunodeficient conditions.
 SQ Sequence 164 AA;
 SQ 6 A; 12 R; 9 N; 5 D; 0 B; 4 C; 8 Q; 12 E; 0 Z; 7 G; 7 H;
 SQ 11 I; 21 L; 10 K; 5 M; 10 F; 2 P; 10 S; 10 T; 2 W; 8 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
RKISLFSCLKDRHDFGFPQEEFGNQFQKAETILTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN
   30      40      50      60      70 X      80      90

HLKTVLEEKLEKEDFTRGKLMSSLHLKRYVGRILHYL
   100     110     120     130

```

12. US-08-249-182-6 (1-9)

P30227 Sequence of hybrid interferon HuFIN-alpha-61A-beta

ID P30227 standard; Protein; 165 AA.
 AC P30227;
 DT 25-MAY-1992 (first entry)
 DE Sequence of hybrid interferon HuFIN-alpha-61A-beta-1
 DE composed of AAs 1-41 of HuIFN-alpha-61A and AAs 43-166 of
 DE HuIFN-beta-1.
 KW Hybrid interferon; antiviral; therapy; cancer; tumour.
 OS Homo sapiens.
 PN W08302461-A.
 PD 21-JUL-1983.
 PF 18-JAN-1983; 900607.
 PR 19-JAN-1982; US-340782.
 PR 03-FEB-1983; US-463574.
 PR 15-JUL-1985; US-755265.
 PA (CETU-) CETUS CORP.
 PI Mark DF, Creasey AA;
 DR WPI; 83-723186/30.
 DR N-PSDB; N30160.
 PT Multi:class hybrid interferon poly:peptide(s) - with restricted
 PT antiviral and cell growth regulatory activities
 PS Example; Fig 23; 61pp; English.
 CC The inventors claim a multiclass hybrid interferon polypeptide and a
 CC DNA unit having a nucleotide sequence which encodes it. Pref. the
 CC AA sequence consists of alpha and beta interferons. Pref. IF1 is
 CC (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
 CC of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
 CC HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
 CC N30160, P30227). Alternatively IF1 is the amino terminal end of a
 CC beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
 CC 1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
 CC resp.) (see N30156, P30223). In the examples plasmids pGW5 and
 CC pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,

10. US-08-249-182-6 (1-9)
P80055 Sequence of human interferon (huIFN) alpha-61A-beta

ID P80055 standard; protein; 161 AA.
AC P80055;
DT 17-NOV-1990 (first entry)
DE Sequence of human interferon (huIFN) alpha-61A-beta-1 hybrid
KW Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW antitumour; antiviral; therapy.
OS Homo sapiens.
PN US4758428-A.
PD 19-JUL-1988.
PF 15-JUL-1985; 755265.
PR 19-JAN-1983; CA-419758.
PR 15-JUL-1985; US-755265.
PA (CETU) Cetus Corp.
PI Mark DF, Creasey AA;
DR WPI; 88-219882/31.
DR N-PSDB; n80052.
PT Multi-class hybrid interferon polypeptide(s) -
PT having sequence from interferon-alpha-1 and sequence from
PT interferon-beta-1 for restricted activity
PS Example; Fig 23; 24pp; English.
CC Multi-class hybrid interferon polypeptides having an AA sequence composed
CC of 2 distinct subsequences are claimed. The plasmids used in the
CC construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC pDM101/trp/beta-1. The hybrid gene was constructed by taking advantage of
CC homologies between huIFN alpha-61A and huIFN beta-1 at around AA 40 of
CC both proteins. The sequence 5'-proximal to the DdeI restriction enzyme
CC cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC 3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC 2 genes while preserving the translational reading frame of both genes.
CC This hybrid interferon is denoted huIFN alpha-61A-beta-1. The amino
CC terminal portion of this polypeptide starting with methionine is composed
CC of the sequence 1-41 of huIFN alpha-61A and the carboxy terminal portion
CC is composed of AAs 47-166 of huIFN beta-1.
SQ Sequence 161 AA;
SQ 7 A; 12 R; 10 N; 6 D; 0 B; 3 C; 9 Q; 10 E; 0 Z; 6 G; 7 H;
SQ 11 I; 20 L; 8 K; 6 M; 9 F; 3 P; 10 S; 9 T; 2 W; 8 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                YDVPWNETI
                                |||||
AAGNGRISPFSCCLKDRHDFGFPQLQGFQKEDAALTIYEMLNIFAIFRQDSSSTGWNETIVENLLANVYHQIN
20          30          40          50          60          70          80          90

HLKTVLEEKLEKEDFTRGKLMSSHLKRYVGRILHYL
100          110          120
```

11. US-08-249-182-6 (1-9)
P50281 Protein sequence encoding modified interferon-beta

ID P50281 standard; protein; 164 AA.
AC P50281;
DT 09-DEC-1991 (first entry)
DE Protein sequence encoding modified interferon-beta.
KW Interferon-beta; IFNX406; virucide; antitumor; immunostimulant.
OS Synthetic.
PN EP-131816-A.
PD 23-JAN-1985.
PF 28-JUN-1984; 107498.

X X
YDVPWNETI
|| |||

PEVESQPLDLSQKKEKQSEYEQQVVKSIKPQKSEPQPYSTYAKAPIWESYDFDWNEDDAKFI LPAPYRLTK
30 40 50 60 70 X 80 90

ADEIVLGSKIVKLRTIIETAIKTONYSALPEAVFELD
100 110 120 130

9. US-08-249-182-6 (1-9)

P60691 Sequence of the region coding for a hybrid protein

ID P60691 standard; protein; 161 AA.
AC P60691;
DT 06-AUG-1991 (first entry)
DE Sequence of the region coding for a hybrid protein
DE comprising human interferon-alpha-61A and interferon-beta-1
DE from plasmid p-alpha-61A and plasmid pDM101/trp/beta-1
KW Plasmid pDM101/trp/beta-1; human; interferon-beta-1; virucide;
KW immunostimulant; plasmid p-alpha-61A; chimeric protein;
KW chimeric gene; hybrid protein; ss.
OS Homo sapiens.
PN US4569908-A.
PD 11-FEB-1986.
PF 03-FEB-1983; 463574.
PR 19-JAN-1982; US-340782.
PR 19-JAN-1983; CA-419758.
PR 03-FEB-1983; US-463574.
PA (CETU) CETUS CORP.
PI Mark DF, Creasey AA;
DR WPI; 86-061770/09.
DR N-PSDB; N60657.
PT DNA producing hybrid interferon polypeptide - which has activity
PT restricted to cell growth regulatory or antiviral activity
PS Disclosure; Fig. 23; 25pp; English.
CC The N-terminal end coding region of the HuIFN-alpha-61A DNA is
CC fused to the DNA coding for the C-terminal end region of
CC HuIFN-beta-1 in such a way that the translational reading
CC frame of the 2 proteins are preserved and the resulting
CC protein being expressed from this hybrid gene will have the
CC AA sequence of HuIFN-alpha-61A at its N-terminal portion and
CC the AA sequence of HuIFN-beta-1 at its C-terminal portion.
CC The chimeric protein may be used to treat tumors and cancers
CC without immunosuppression side effects. They may also be used
CC to treat encephalomyocarditis virus infection, rabies and other
CC viral zoonoses and arbo virus infections. See also N60652-6
CC and P60686-90.
SQ Sequence 161 AA;
SQ 7 A; 12 R; 9 N; 6 D; 0 B; 3 C; 9 Q; 10 E; 0 Z; 6 G; 7 H;
SQ 11 I; 20 L; 9 K; 6 M; 9 F; 3 P; 10 S; 9 T; 2 W; 8 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
|||||

AQMGRI SPFSCLKDRHDFGFPQLQQFQKEDAALTIYEMLQNI FAIFRQDSSTGWN ETIVENLLANVYHQIN
20 30 40 50 60 70 80 90

HLKTVLEEKLEKEDFTRGKLMSLHLKRY YGRILHYL
100 110 120

CC the vector. (D) Prodn. of a yeast or non-yeast polypeptide or its
CC deriv. by culturing the transformed yeast. The polypeptides are
CC hormones, antiviral and anticancer peptides, enzymes and interferon.
SQ Sequence 155 AA;
SQ 6 A; 10 R; 11 N; 5 D; 0 B; 3 C; 10 Q; 13 E; 0 Z; 5 G; 5 H;
SQ 11 I; 21 L; 11 K; 3 M; 8 F; 1 P; 8 S; 7 T; 3 W; 9 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
|||||
NGRLEYCLKDRMNFDIPEEIKQLQGFQKEDAALTIYEMLNIFAIFRQSSSTGWNETIVENLLANVYHQIN
20 30 40 50 60 X 70 X 80
HLKTVLEEKLEKEDFTRGKLMSSLHLKRYVGRILHYL
90 100 110 120

8. US-08-249-182-6 (1-9)

R29179 Astrovirus serotype A2 clone A43351-encoded polype

ID R29179 standard; Protein; 156 AA.
AC R29179;
DT 08-APR-1993 (first entry)
DE Astrovirus serotype A2 clone A43351-encoded polypeptide.
KW Sequence-independent single primer amplification; SISPA;
KW viral gastroenteritis.
OS Astrovirus.
PN W09220803-A.
PD 26-NOV-1992.
PF 20-MAY-1992; U04276.
PR 20-MAY-1991; US-702731.
PA (GENE-) GENELABS INC.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Greenberg HB, Kim JP, Matsui SM, Reyes GR;
DR WPI; 92-415781/50.
DR N-PSDB; 031811.
PT Recombinant Astrovirus polynucleotide(s) (e.g. cDNA) - and
PT corresponding polypeptide antigens and antibodies, useful in
PT diagnosis of Astrovirus infection and as vaccines
PS Claim 3; Fig 5; 85pp; English.
CC The cell lines A1 LLCMK2 and LLCMK2 (i.e. monkey kidney cell lines
CC infected with Astrovirus type 1 and uninfected cells, respectively)
CC were cultured and RNA was extracted. The RNA was used to generate
CC cDNA which was amplified by the SISPA method. The amplified cDNA was
CC cloned into lambda gt11, packaged and used to infect E.coli KM392.
CC The cDNA library was screened using astrovirus-immunised rabbit sera
CC to obtain ten clones containing Astroviral coding sequences (i.e.
CC clones A35, A43, A39, A1, A2, A11, A13, A14, A21 and A33). The
CC clones can be used as probes to obtain a complete set of
CC overlapping genomic cDNA clones. The sequence of A43351 is made up
CC of the overlapping clones A43, A35 and A1. The sequence of the
CC A43351 polypeptide referred to as SEQ ID NO 14 in the claims
CC consists of amino acids 7 to 150.
CC See 031806-031816 and 037634.
SQ Sequence 156 AA;
SQ 15 A; 6 R; 2 N; 7 D; 0 B; 1 C; 10 Q; 17 E; 0 Z; 4 G; 0 H;
SQ 9 I; 11 L; 19 K; 0 M; 6 F; 13 P; 11 S; 9 T; 2 W; 7 Y; 7 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

FN EP--76487-A.
 PD 13-APR-1983.
 PF 01-SEP-1982; 024871.
 PR 03-OCT-1981; GB-029937.
 PR 26-MAR-1982; GB-008988.
 PR 01-SEP-1982; GB-024871.
 PA (CIBA) CIBA GEIGY AG.
 PI Meyer F;
 DR WPI; 83-37194K/16.
 DR N-PSDB; N30101.
 PT Human interferon polypeptide - produced by host cells, esp.
 PT E.coli transformed with recombinant DNA contg. sequence
 PT obtainable from human lymphoblastoid cells
 PS Claim 42; Page 108-9 and Fig 5; 142pp; German.
 CC The inventors claim DNA encoding a polypeptide with human
 CC lymphoblastoid interferon activity, hosts which have been
 CC transformed with the DNA and the polypeptide. Variants of the DNA
 CC sequence are also claimed. The polypeptides are useful as immuno-
 CC modulators, esp. as antiviral, antitumour and anticancer agents.
 SQ Sequence 155 AA;
 SQ 6 A; 10 R; 11 N; 5 D; 0 B; 3 C; 10 Q; 13 E; 0 Z; 5 G; 5 H;
 SQ 11 I; 21 L; 11 K; 3 M; 8 F; 1 P; 8 S; 7 T; 3 W; 9 Y; 5 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 |||||

NGRLEYCLKDRMNFDIPEEIKQLQGFQKEDAALTIYEMLNIFAIFRQDSSTGWNETIVENLLANVYHQIN
 20 30 40 50 60 X 70 X 80

HLKTVLEEKLEKEDFTRGKLMSSLHLKRYVGRILHYL
 90 100 110 120

7. US-08-249-182-6 (1-9)
 P40124 Sequence encoded by the cDNA insert of the recomb

ID P40124 standard; Protein; 155 AA.
 AC P40124;
 DT 12-FEB-1992 (first entry)
 DE Sequence encoded by the cDNA insert of the recombinant plasmid
 DE CG-pBR 322/HLycIFN-beta1.
 KW Yeast expression vector; promoter.
 OS Homo sapiens.
 PN EP-100561-A.
 PD 15-FEB-1984.
 PF 08-AUG-1983; 107804.
 PR 09-AUG-1982; GB-022883.
 PR 31-DEC-1982; GB-037026.
 PR 02-JUN-1983; GB-015145.
 PR 14-JUL-1983; GB-019099.
 PA (CIBA) CIBA GEIGY AG.
 PI Hinnen A, Meyhack B, Meyer F;
 DR WPI; 84-044243/08.
 DR N-PSDB; N40108.
 PT Acid phosphatase promoting DNA fragment - for expressing
 PT peptide(s) in yeasts which are more easily cultured than e.coli
 PS Example; Fig 11; 166pp; English.
 CC The inventors claim: (A) DNA fragment consisting of yeast acid
 CC phosphatase promoter and flanking sequences or its mutants which
 CC retain the promoter function. (B) Hybrid vector consisting of a
 CC yeast acid phosphatase promoter and a yeast or non-yeast polypeptide
 CC coding region controlled by the promoter. (C) A yeast transformed by

5. US-08-249-182-6 (1-9)

R29174 Astrovirus serotype A1 clone A35-encoded polypepti

ID R29174 standard; Protein; 75 AA.
AC R29174;
DT 08-APR-1993 (first entry)
DE Astrovirus serotype A1 clone A35-encoded polypeptide.
KW Sequence-independent single primer amplification; SISPA;
KW viral gastroenteritis.
OS Astrovirus.
PN W09220803-A.
PD 26-NOV-1992.
PF 20-MAY-1992; U04276.
PR 20-MAY-1991; US-702731.
PA (GENE-) GENELABS INC.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Greenberg HB, Kim JP, Matsui SM, Reyes GR;
DR WPI; 92-415781/50.
DR N-PSDB; Q31806.
PT Recombinant Astrovirus polynucleotide(s) (e.g. cDNA) - and
PT corresponding polypeptide antigens and antibodies, useful in
PT diagnosis of Astrovirus infection and as vaccines
PS Claim 3; Page 50-51; 85pp; English.
CC The cell lines A1 LLCMK2 and LLCMK2 (i.e. monkey kidney cell lines
CC infected with Astrovirus type 1 and uninfected cells, respectively)
CC were cultured and RNA was extracted. The RNA was used to generate
CC cDNA which was amplified by the SISPA method. The amplified cDNA was
CC cloned into lambda gt11, packaged and used to infect E.coli KM392.
CC The cDNA library was screened using astrovirus-immunised rabbit sera
CC to obtain ten clones containing Astroviral coding sequences (i.e.
CC clones A35, A43, A39, A1, A2, A11, A13, A14, A21 and A33). The
CC clones can be used as probes to obtain a complete set of
CC overlapping genomic cDNA clones.
CC See also Q31807-Q31816 and Q37634.
SQ Sequence 75 AA;
SQ 5 A; 2 R; 1 N; 5 D; 0 B; 0 C; 5 Q; 7 E; 0 Z; 2 G; 0 H;
SQ 7 I; 4 L; 8 K; 0 M; 2 F; 6 P; 6 S; 4 T; 2 W; 5 Y; 4 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
Residue Identity = 55% Matches = 5 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
YDVPWNETI
|| ||
SEYEQQVVKSIKPKSEPPYSQTYGKAPIWESYDFDWNEDDAKFI LPAPYRLTKADEIVLGSKIVKLRTII
10 20 30 X 40 X 50 60 70

ETA

6. US-08-249-182-6 (1-9)

P30180 Sequence of a polypeptide with human lymphoblastoi

ID P30180 standard; Protein; 155 AA.
AC P30180;
DT 14-JUN-1992 (first entry)
DE Sequence of a polypeptide with human lymphoblastoid interferon
DE activity encoded by plasmid CG-pBR 322/HL gamma cIFN-beta1.
KW Immunomodulator; antiviral agent; antitumor; anticancer.
OS Homo sapiens.

CC ced-4 sequences. These peptides were used to construct degenerate
 CC primers and probes.
 SQ Sequence 30 AA;
 SQ 1 A; 4 R; 1 N; 3 D; 0 B; 1 C; 2 Q; 4 E; 0 Z; 0 G; 0 H;
 SQ 2 I; 3 L; 0 K; 0 M; 0 F; 0 P; 1 S; 3 T; 1 W; 0 Y; 4 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 || ||
 DDVVQEEETIRWAGELRLRCLVTTRDVEISN
 X 10 20 30

4. US-08-249-182-6 (1-9)
 R27732 SalF20.5R.

ID R27732 standard; peptide; 59 AA.
 AC R27732;
 DT 09-MAR-1993 (first entry)
 DE SalF20.5R.
 KW Virus vector; vaccinia virus; papillomavirus; HPV;
 KW immunotherapeutic.
 OS Vaccinia virus.
 PN W09216636-A.
 PD 01-OCT-1992.
 PF 10-MAR-1992; G00424.
 PR 14-MAR-1991; GB-005383.
 PA (IMMU) IMMUNOLOGY LTD.
 PI Boursnell MEG, Inglis SC, Munro AJ;
 DR WPI; 92-349219/42.
 DR N-PDSB; Q29392.
 PT Recombinant virus vectors encoding human papillomavirus proteins
 PT - for treating and vaccinating against HPV infections and
 PT conditions caused by them, such as cervical cancer
 PS Disclosure; Fig 10; 83pp; English.
 CC To make a recombinant virus vector comprising human papillomavirus
 CC genes inserted into the vaccinia virus genome, neutral sites
 CC for insertion must be utilised such that replicative ability is not
 CC adversely affected. The neutral sites are identified by analysing
 CC the viral genome to identify ORFs which are likely to encode
 CC functional genes and selecting sites between such ORFs or within
 CC sequences for non-functional genes. One such neutral site is site B,
 CC present in an intergenic region between SalF20R and SalF20.5R. It is
 CC placed 70 bases upstream of SalF20.5R to avoid promoter elements
 CC associated with that gene. Also there is no transcription
 CC termination signal with which site B could interfere and hence the
 CC sequence is suitable as a neutral insertion site. HPV DNA sequences
 CC may be inserted at this site, e.g. those encoding E6 or E7 of HPV 16
 CC and 18 or mutants of these proteins. The recombinant virus vector may
 CC be used immunotherapeutically to activate cells of the immune system
 CC against HPV. See also R27723-43.
 SQ Sequence 59 AA;
 SQ 3 A; 1 R; 7 N; 4 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 3 H;
 SQ 8 I; 6 L; 1 K; 1 M; 3 F; 0 P; 1 S; 4 T; 1 W; 6 Y; 9 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 YDVPWNETI
 || ||

ID R42781 standard; Protein; 30 AA.
 AC R42781;
 DT 28-APR-1994 (first entry)
 DE Consensus peptide, (ced-4)
 KW Long-distance homology; evolution; nematode;
 KW hybridisation; lower organism; structural homologue;
 KW Alzheimer's disease; cell death gene; PCR; polymerase chain reaction;
 KW ciona intestinalis; echinoderm; lamprey; puffer fish;
 KW mammal; probe.
 OS Synthetic.
 PN W09320237-A.
 PD 14-OCT-1993.
 PF 01-APR-1993; U03102.
 PR 01-APR-1992; US-861458.
 PA (CAMB-) CAMBRIDGE NEUROSCIENCE INC.
 PI Johnson CD, Marchionni MA;
 DR WPI; 93-336943/42.
 PT Long-distance homology cloning of genes from lower organisms -
 PT used to identify DNA that codes for evolutionary conserved
 PT aminoacid sequences
 PS Disclosure; Fig 25; 188pp; English.
 CC Sequences (R42771-87) show conserved stretches within deduced
 CC ced-4 sequences. These peptides were used to construct degenerate
 CC primers and probes.
 SQ Sequence 30 AA;
 SQ 1 A; 4 R; 1 N; 4 D; 0 B; 2 C; 2 Q; 3 E; 0 Z; 0 G; 0 H;
 SQ 3 I; 3 L; 0 K; 0 M; 0 F; 0 P; 0 S; 3 T; 1 W; 0 Y; 3 V;

Initial Score = 5 Optimized Score = 5 Significance = 4.34
 Residue Identity = 55% Matches = 5 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

X      X
YDVPWNETI
  ||  |||
DDVVQDETIRWAGELRLRLCLITTRDVEICN
X      10      20      30
    
```

3. US-08-249-182-6 (1-9)

R42780 Consensus peptide, (ced-4)

ID R42780 standard; Protein; 30 AA.
 AC R42780;
 DT 28-APR-1994 (first entry)
 DE Consensus peptide, (ced-4)
 KW Long-distance homology; evolution; nematode;
 KW hybridisation; lower organism; structural homologue;
 KW Alzheimer's disease; cell death gene; PCR; polymerase chain reaction;
 KW ciona intestinalis; echinoderm; lamprey; puffer fish;
 KW mammal; probe.
 OS Synthetic.
 PN W09320237-A.
 PD 14-OCT-1993.
 PF 01-APR-1993; U03102.
 PR 01-APR-1992; US-861458.
 PA (CAMB-) CAMBRIDGE NEUROSCIENCE INC.
 PI Johnson CD, Marchionni MA;
 DR WPI; 93-336943/42.
 PT Long-distance homology cloning of genes from lower organisms -
 PT used to identify DNA that codes for evolutionary conserved
 PT aminoacid sequences
 PS Disclosure; Fig 25; 188pp; English.
 CC Sequences (R42771-87) show conserved stretches within deduced

22. P50196	Modified human interferon (IF	166	5	5	4.34	0
23. P50195	Modified human interferon (IF	166	5	5	4.34	0
24. P50194	Modified human interferon (IF	166	5	5	4.34	0
25. P50193	Modified human interferon (IF	166	5	5	4.34	0
26. P50192	Modified human interferon (IF	166	5	5	4.34	0
27. P50279	Protein sequence encoding syn	166	5	5	4.34	0
28. P50278	Protein sequence encoding mod	166	5	5	4.34	0
29. P50277	Protein sequence encoding mod	166	5	5	4.34	0
30. P50276	Protein sequence encoding mod	166	5	5	4.34	0
31. P50275	Protein sequence encoding mod	166	5	5	4.34	0
32. P50274	Protein sequence encoding mod	166	5	5	4.34	0
33. P50272	Protein sequence encoding mod	166	5	5	4.34	0
34. P50271	Protein sequence encoding mod	166	5	5	4.34	0
35. P50270	Protein sequence encoding mod	166	5	5	4.34	0
36. P50262	Sequence encoded by the seque	166	5	5	4.34	0
37. P50591	Sequence of the interferon mu	166	5	5	4.34	0
38. P50032	Sequence of new modified huma	166	5	5	4.34	0
39. P50031	Sequence of new modified huma	166	5	5	4.34	0
40. P50030	Sequence of new modified huma	166	5	5	4.34	0

1. US-08-249-182-6 (1-9)

R37448 Autotaxin peptide ATX 47.

ID R37448 standard; peptide; 9 AA.
AC R37448;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 47.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 47. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 9 AA;
SQ 0 A; 0 R; 1 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 0 G; 0 H;
SQ 1 I; 0 L; 0 K; 0 M; 0 F; 1 P; 0 S; 1 T; 1 W; 1 Y; 1 V;

Initial Score = 9 Optimized Score = 9 Significance = 7.81
Residue Identity = 100% Matches = 9 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

X      X
YDVPWNETI
|||||||
YDVPWNETI
X      X

```

2. US-08-249-182-6 (1-9)

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	0	1	1.15

Times:	CPU	Total Elapsed
	00:00:27.93	00:00:30.00

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	3921

Cut-off raised to 2.

Cut-off raised to 3.

Cut-off raised to 4.

The scores below are sorted by initial score.

Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.			Sig.	Frame
		Length	Score	Score		
1. R37448	Autotaxin peptide ATX 47.	9	9	9	7.81	0

The list of other best scores is:

Sequence Name	Description	Init. Opt.			Sig.	Frame
		Length	Score	Score		

**** 4 standard deviations above mean ****

2. R42781	Consensus peptide, (ced-4)	30	5	5	4.34	0
3. R42780	Consensus peptide, (ced-4)	30	5	5	4.34	0
4. R27732	SalF20.5R.	59	5	5	4.34	0
5. R29174	Astrovirus serotype A1 clone	75	5	5	4.34	0
6. P30180	Sequence of a polypeptide wit	155	5	5	4.34	0
7. P40124	Sequence encoded by the cDNA	155	5	5	4.34	0
8. R29179	Astrovirus serotype A2 clone	156	5	5	4.34	0
9. P60691	Sequence of the region coding	161	5	5	4.34	0
10. P80055	Sequence of human interferon	161	5	5	4.34	0
11. P50281	Protein sequence encoding mod	164	5	5	4.34	0
12. P30227	Sequence of hybrid interferon	165	5	5	4.34	0
13. P50207	Modified human interferon (IF	165	5	5	4.34	0
14. P50280	Protein sequence encoding mod	165	5	5	4.34	0
15. P50273	Protein sequence encoding mod	165	5	5	4.34	0
16. P20025	Sequence encoded by human fib	166	5	5	4.34	0
17. P30222	Sequence of hybrid interferon	166	5	5	4.34	0
18. P30219	Sequence of interferon (HuIFN	166	5	5	4.34	0
19. P40626	Sequence of interferon-beta-S	166	5	5	4.34	0
20. P50206	Modified human interferon (IF	166	5	5	4.34	0
21. P50197	Modified human interferon (IF	166	5	5	4.34	0

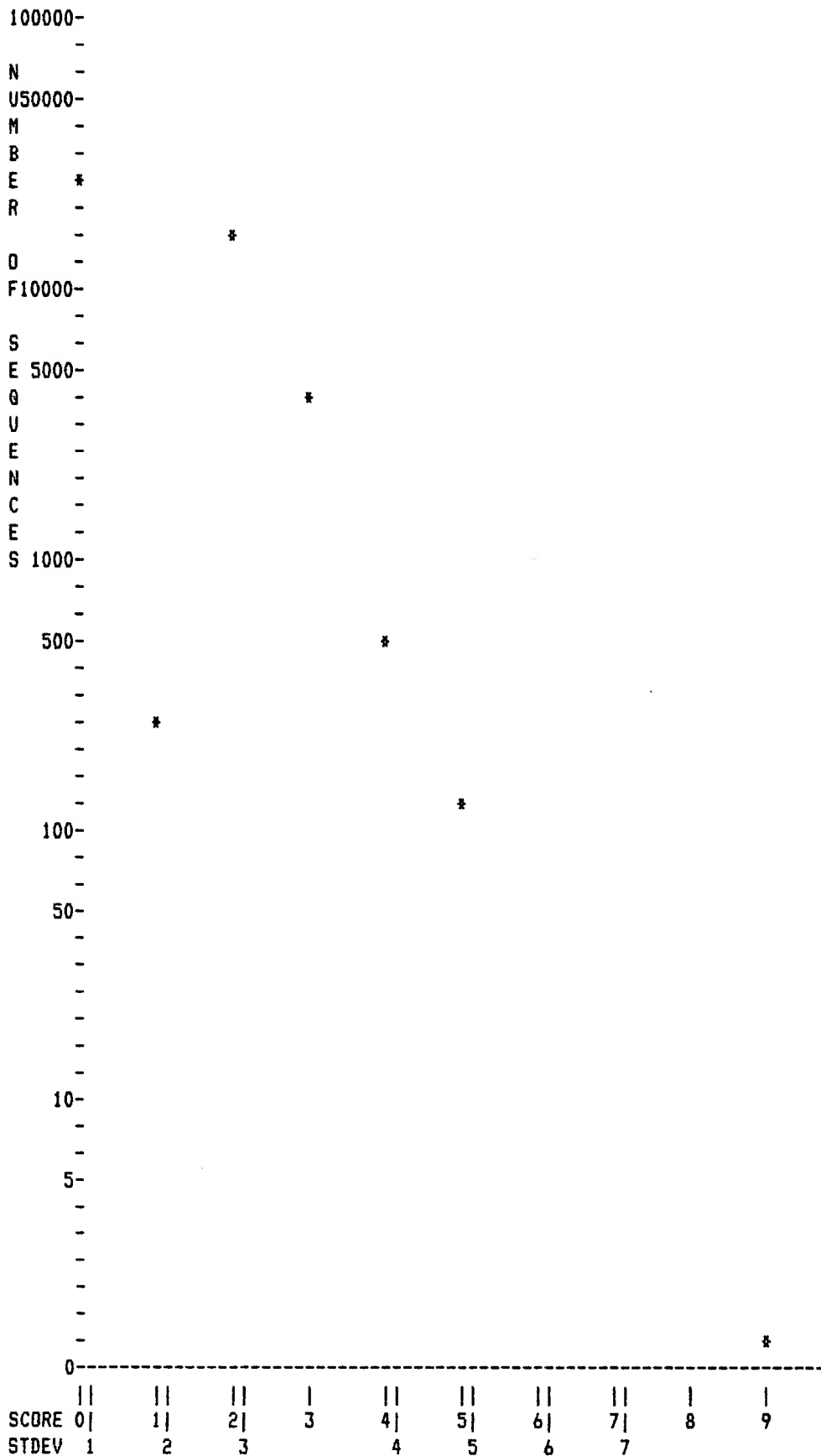
Results file u249_6a.res made by on Thu 22 Sep 94 10:51:41-PDT.

Query sequence being compared:US-08-249-182-6 (1-9)

Number of sequences searched: 42145

Number of scores above cutoff: 3921

Results of the initial comparison of US-08-249-182-6 (1-9) with:
Data bank : A-GeneSeq 15, all entries



FT DISULFID 34 37 BY SIMILARITY.
SQ SEQUENCE 61 AA; 7033 MW; 16901 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.49
Residue Identity = 50% Matches = 5 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                X      X
                PEEVTRPNYL
                ||  |  ||
FTCFTTPSDTSETCPDGGNICYEKRWNSHQVEIKGCVASCFEFESRFRYLLCCRIDNCNK
      10      20      30      40 X      50      60
```

15. US-08-249-182-5 (1-10)

NUCC_SECCE NADH-PLASTOQUINONE OXIDOREDUCTASE 49 KD SUBUNIT, C

ID NUCC_SECCE STANDARD; PRT; 90 AA.
AC P27758;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE NADH-PLASTOQUINONE OXIDOREDUCTASE 49 KD SUBUNIT, CHLOROPLAST
DE (EC 1.6.5.3) (FRAGMENT).
GN NDHH.
OS SECALE CEREALE (RYE).
OG CHLOROPLAST.
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
OC CYPERALES; GRAMINEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 90037033
RA PROMBONA A., SUBRAMANIAN A.R.;
RL J. BIOL. CHEM. 264:19060-19065(1989).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE PLASTOQUINONE. COMPONENT OF THE IRON-SULFUR (IP) FRAGMENT OF
CC THE ENZYME.
CC -!- CATALYTIC ACTIVITY: NADH + PLASTOQUINONE = NAD(+) + PLASTOQUINOL.
CC -!- SIMILARITY: TO OTHER COMPLEX I 49 KD SUBUNITS AND TO E.COLI
CC FORMATE HYDROGENLYASE (FHL) SUBUNIT HYCE.
DR EMBL; X14557; CHSCRPS3.
DR PIR; C34435; C34435.
DR PROSITE; PS00535; COMPLEX1_49K.
KW OXIDOREDUCTASE; NAD; PLASTOQUINONE; CHLOROPLAST; IRON-SULFUR; 4FE-4S.
FT NON_TER 1 1
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 10509 MW; 42181 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.49
Residue Identity = 50% Matches = 5 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                X      10
                PEEVTRPNYL
                |||  ||
IVTLDCGEDVIDCEPILGYLHRGMEKIAENRTIIQYLPYVTRWDYLATMFTEAITVNAPEFLENIQIPQRASY
      10      20      30      X 40      X 50      60      70

IRVIMLELSRIASHLLWL
      80      90
```

> 0 <
0| 10 IntelliGenetics
> 0 < *Seq. 6*

ID CRYA_BACTIN STANDARD; PRT; 1228 AA.
 AC P05517;
 DT 01-NOV-1988 (REL. 09, CREATED)
 DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
 DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
 DE 139 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL
 DE PROTOXIN).
 GN CRYA4.
 OS BACILLUS THURINGIENSIS (SUBSP. KURSTAKI).
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-2;
 RM 88203216
 RA BRIZZARD B.L., WHITELEY H.R.;
 RL NUCLEIC ACIDS RES. 16:2723-2723(1988).
 CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF INSECTS.
 CC -!- TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-TERMINUS.
 CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 DR EMBL; X06711; BTCRYA4.
 DR PIR; S00873; S00873.
 KW TOXIN; SPORULATION.
 SQ SEQUENCE 1228 AA; 139629 MW; 7697478 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 PEEVTRPNYL
 || ||||
 GIYLEPIHGVPTVRFNFTNPQNISDRGTANYSQPYESPGLOLKDSETELPPEPPERPNYESYSHRLSHIGII
 400 410 420 430 440 450 460 470
 LQSRVNVPPVYSWTHRSADRTNTIGPNRITQIPMVKASE
 480 490 500

14. US-08-249-182-5 (1-10)
 TXW2_NAJHH WEAK TOXIN CM-2.

ID TXW2_NAJHH STANDARD; PRT; 61 AA.
 AC P01415;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE WEAK TOXIN CM-2.
 OS NAJA HAJE HAJE (EGYPTIAN COBRA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=VENOM;
 RM 79045337
 RA JOUBERT F.J., TALJAARD N.;
 RL EUR. J. BIOCHEM. 90:359-367(1978).
 CC -!- LD(50) IS 16.1 MG/KG BY SUBCUTANEOUS INJECTION.
 DR PIR; A01685; T2NJ2Y.
 DR PROSITE; PS00272; SNAKE_TOXIN.
 KW VENOM; TOXIN.
 FT DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 14 37 BY SIMILARITY.
 FT DISULFID 41 53 BY SIMILARITY.

12. US-08-249-182-5 (1-10)

K6PF_RABIT 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (

ID K6PF_RABIT STANDARD; PRT; 779 AA.
 AC P00511;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE 6-PHOSPHOFRUCTOKINASE, MUSCLE TYPE (EC 2.7.1.11) (PHOSPHOFRUCTOKINASE
 DE 1) (PHOSPHOHEXOKINASE).
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; LAGOMORPHA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE;
 RM 87166033
 RA LEE C.P., KAO M.C., FRENCH B.A., PUTNEY S.D., CHANG S.H.;
 RL J. BIOL. CHEM. 262:4195-4199(1987).
 RN [2]
 RP SEQUENCE.
 RM 84219739
 RA POORMAN R.A., RANDOLPH A., KEMP R.G., HEINRIKSON R.L.;
 RL NATURE 309:467-469(1984).
 CC -!- CATALYTIC ACTIVITY: ATP + D-FRUCTOSE 6-PHOSPHATE = ADP +
 CC D-FRUCTOSE 1,6-BISPHOSPHATE.
 CC -!- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- COFACTOR: REQUIRES MAGNESIUM ION.
 CC -!- ENZYME REGULATION: ALLOSTERIC ENZYME ACTIVATED BY ADP, AMP, OR
 CC FRUCTOSE BISPHOSPHATE AND INHIBITED BY ATP OR CITRATE.
 DR PIR; A00604; KIRBF.
 DR PIR; A26550; A26550.
 DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE.
 KW KINASE; TRANSFERASE; GLYCOLYSIS; DUPLICATION; ALLOSTERIC ENZYME;
 KW PHOSPHORYLATION; MAGNESIUM; MULTIGENE FAMILY.
 FT INIT_MET 0 0
 FT REPEAT 1 401 APPROXIMATE.
 FT REPEAT 402 779 APPROXIMATE.
 FT MOD_RES 774 774 PHOSPHORYLATION.
 FT CONFLICT 268 268 R -> S (IN REF. 2).
 FT CONFLICT 479 507 MISSING (IN REF. 2).
 FT CONFLICT 558 558 S -> I (IN REF. 2).
 FT CONFLICT 565 565 MISSING (IN REF. 2).
 SQ SEQUENCE 779 AA; 85072 MW; 3033380 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 PEEVTRPNYL
 || || ||
 GRLRAAHNLVKRGITNLCVIGDGLTGADTFRSEWSDLLSLQKAGKITAEATRSSLNIVGLVGSIDND
 100 110 120 130 140 X 150 X 160
 FCGTDMTIGTDSALHRITEIVDAITTTAGSHQRTFVLE
 170 180 190 200

13. US-08-249-182-5 (1-10)

CRYX_BACTK 139 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALI

DR PROSITE; PS00727; AP_NUCLEASE_F1_1.
 DR PROSITE; PS00728; AP_NUCLEASE_F1_3.
 KW DNA REPAIR; ENDDNUCLEASE; HYDROLASE; NUCLEASE; NUCLEAR PROTEIN.
 FT DOMAIN 428 679 AP ENDDNUCLEASE.
 SQ SEQUENCE 679 AA; 74662 MW; 2298057 CN;

Initial Score = 6 Optimized Score = 7 Significance = 4.65
 Residue Identity = 72% Matches = 8 Mismatches = 2
 Gaps = 1 Conservative Substitutions = 0

X X
 PEEVTR-PNYL
 ||||| |
 PADKEFNLIKCSWNVAGLRAWLKKDGLQLIDLEEDIFCLQETKCANDQLPEEVTRLPGYHPYWLCPGGYA
 420 430 440 450 460 470 480 490

GVAIYSKIMPIHVEYGIGNEEFDDVGRMITAEYEKFYLI
 500 510 520 530

10. US-08-249-182-5 (1-10)
 G13A_DICD1 CELL SURFACE GLYCOPROTEIN GP138A PRECURSOR.

ID G13A_DICD1 STANDARD; PRT; 730 AA.
 AC P34115;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE CELL SURFACE GLYCOPROTEIN GP138A PRECURSOR.
 GN FUSA.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;
 OC EUMYCETOZOA; DICTYOSTELIA.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RM 93193972
 RA FANG H., HIGA M., SUZUKI K., AIBA K., URUSHIHARA H., YANAGISAWA K.;
 RL DEV. BIOL. 156:201-208(1993).
 CC -!- FUNCTION: INVOLVED IN THE SEXUAL CELL FUSION OF D.DISCOIDEUM.
 DR EMBL; D12883; DDGP138A.
 DR DICTYDB; DD03014; FUSA.
 KW GLYCOPROTEIN; SIGNAL; MULTIGENE FAMILY.
 FT SIGNAL 1 20
 FT CHAIN 21 730 CELL SURFACE GLYCOPROTEIN GP138A.
 FT CARBOHYD 58 58 POTENTIAL.
 FT CARBOHYD 89 89 POTENTIAL.
 FT CARBOHYD 124 124 POTENTIAL.
 FT CARBOHYD 198 198 POTENTIAL.
 FT CARBOHYD 224 224 POTENTIAL.
 FT CARBOHYD 392 392 POTENTIAL.
 FT CARBOHYD 420 420 POTENTIAL.
 FT CARBOHYD 435 435 POTENTIAL.
 FT CARBOHYD 482 482 POTENTIAL.
 FT CARBOHYD 498 498 POTENTIAL.
 FT CARBOHYD 523 523 POTENTIAL.
 FT CARBOHYD 534 534 POTENTIAL.
 FT CARBOHYD 596 596 POTENTIAL.
 FT CARBOHYD 605 605 POTENTIAL.
 FT CARBOHYD 614 614 POTENTIAL.
 FT CARBOHYD 620 620 POTENTIAL.
 FT CARBOHYD 621 621 POTENTIAL.
 FT CARBOHYD 630 630 POTENTIAL.
 SQ SEQUENCE 730 AA; 80960 MW; 3075521 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65

Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

X      10
PEEVTRPNYL
  || | |||
NTFSYECIDFSNIADFSYYNDYEQYLPNIDNAPLLTGVDIYQSVVVGDIPESYCRINYLSLYYNQLNGTVP
 330      340      350      360      370      X 380      X 390

SCIQCLGGVKGGDIVLPNPFNFNKTTEPYCPTFKIDE
 400      410      420      430
```

11. US-08-249-182-5 (1-10)

G13B_DICDI CELL SURFACE GLYCOPROTEIN GP138B PRECURSOR.

ID G13B_DICDI STANDARD; PRT; 734 AA.
AC P34116;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE CELL SURFACE GLYCOPROTEIN GP138B PRECURSOR.
GN FUSB.
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;
OC EUMYCETOZOA; DICTYOSTELIA.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RM 93193972
RA FANG H., HIGA M., SUZUKI K., AIBA K., URUSHIHARA H., YANAGISAWA K.;
RL DEV. BIOL. 156:201-208(1993).
CC -!- FUNCTION: INVOLVED IN THE SEXUAL CELL FUSION OF D.DISCOIDEUM.
DR EMBL; D12884; DDGP138B.
DR DICTYDB; DD03015; FUSB.
KW GLYCOPROTEIN; SIGNAL; MULTIGENE FAMILY.
FT SIGNAL 1 20
FT CHAIN 21 734 CELL SURFACE GLYCOPROTEIN GP138B.
FT CARBOHYD 58 58 POTENTIAL.
FT CARBOHYD 89 89 POTENTIAL.
FT CARBOHYD 124 124 POTENTIAL.
FT CARBOHYD 198 198 POTENTIAL.
FT CARBOHYD 224 224 POTENTIAL.
FT CARBOHYD 392 392 POTENTIAL.
FT CARBOHYD 420 420 POTENTIAL.
FT CARBOHYD 435 435 POTENTIAL.
FT CARBOHYD 482 482 POTENTIAL.
FT CARBOHYD 498 498 POTENTIAL.
FT CARBOHYD 523 523 POTENTIAL.
FT CARBOHYD 596 596 POTENTIAL.
FT CARBOHYD 605 605 POTENTIAL.
FT CARBOHYD 614 614 POTENTIAL.
FT CARBOHYD 621 621 POTENTIAL.
FT CARBOHYD 630 630 POTENTIAL.
SQ SEQUENCE 734 AA; 81063 MW; 3108617 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.65
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

X      10
PEEVTRPNYL
  || | |||
NTFSYACIDFSNLAYFYDYNEYEQYLPNIDNAPLLNEIYISESVVVGDIPESYCRINYLGGLNYNQLNGTAP
 330      340      350      360      370      X 380      X 390

SCILCLGGNRGGDIVLPNPLNFNKTSEPYCPTFKIDQ
```

ID R29583 standard; Protein; 513 AA.
AC R29583;
DT 19-APR-1993 (first entry)
DE Human activin receptor.
KW Activin receptor; mouse; Xenopus; human; extracellular; ligand binding;
KW hydrophobic; trans-membrane; intracellular; receptor; domain;
KW serine kinase-like; activity; probe; superfamily; secretion signal;
KW golgi membrane; diagnosis; treatment; activin-dependent tumour; brain;
KW neuron; abortion; twinning; wound healing; TGF-beta; immune response;
KW liver regeneration.
OS Homo sapiens.
PN W09220793-A.
PD 26-NOV-1992.
PF 08-MAY-1992; U03825.
PR 10-MAY-1991; US-698709.
PR 09-OCT-1991; US-773229.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
PI Mathews LS, Vale WW;
DR WPI; 92-415771/50.
DR N-PSDB; Q31912.
PT New member of activin-transforming growth factor beta
PT super-family - for diagnosis and treatment of cancer and
PT disorders of the immune, reproductive or central nervous system
PS Disclosure; Page 40; 68pp; English.
CC The sequences given in R29581-83 represent activin receptors from
CC mouse, Xenopus and human respectively. Each of these proteins
CC comprise three distinct domains; an extracellular, ligand binding
CC domain, a hydrophobic, trans-membrane domain and an intracellular,
CC receptor domain having serine kinase-like activity. The DNA sequences
CC encoding these proteins can be used as probes for the identification
CC of additional members of this superfamily of receptor molecules. The
CC proteins may further comprise a second hydrophobic domain at the amino
CC terminal which comprises a secretion signal sequence which promotes
CC the intracellular transport of the initially expressed receptor
CC protein across the golgi membrane. These receptor proteins can be
CC used to develop agents for the diagnosis and/or treatment of eg.
CC activin-dependent tumours, for enhancing the survival of brain
CC neurons, for inducing abortion or twinning in livestock, for
CC stimulating wound healing, for suppression of growth of TGF-beta
CC sensitive tumours, for suppressing immune response, for promoting
CC liver regeneration and for stimulating some immune responses.
SQ Sequence 513 AA;
SQ 38 A; 21 R; 20 N; 28 D; 0 B; 19 C; 18 Q; 37 E; 0 Z; 32 G; 15 H;
SQ 28 I; 46 L; 32 K; 16 M; 20 F; 30 P; 25 S; 24 T; 11 W; 17 Y; 36 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                || || ||
INCYDRTDCVEKKDSPEVVFCCCEGNMCNEKFSYFPMEVETQPTSNPVTPKPPYYNILLYSLVPLMLIAGIV
    90      100      110      120      130 X   140 X   150

ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
    160      170      180      190

```

5. US-08-249-182-7 (1-10)
R29581 Mouse activin receptor.

ID R29581 standard; Protein; 513 AA.
AC R29581;
DT 19-APR-1993 (first entry)

DE mouse activin receptor.
 KW Activin receptor; mouse; Xenopus; human; extracellular; ligand binding;
 KW hydrophobic; trans-membrane; intracellular; receptor; domain;
 KW serine kinase-like; activity; probe; superfamily; secretion signal;
 KW golgi membrane; diagnosis; treatment; activin-dependent tumour; brain;
 KW neuron; abortion; twinning; wound healing; TGF-beta; immune response;
 KW liver regeneration.
 OS Mus musculus.
 PN W09220793-A.
 PD 26-NOV-1992.
 PF 08-MAY-1992; U03825.
 PR 10-MAY-1991; US-698709.
 PR 09-OCT-1991; US-773229.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Mathews LS, Vale WW;
 DR WPI; 92-415771/50.
 DR N-PSDB; 031910.
 PT New member of activin-transforming growth factor beta
 PT super-family - for diagnosis and treatment of cancer and
 PT disorders of the immune, reproductive or central nervous system
 PS Disclosure; Page 42-45; 68pp; English.
 CC The sequences given in R29581-83 represent activin receptors from
 CC mouse, Xenopus and human respectively. Each of these proteins
 CC comprise three distinct domains; an extracellular, ligand binding
 CC domain, a hydrophobic, trans-membrane domain and an intracellular,
 CC receptor domain having serine kinase-like activity. The DNA sequences
 CC encoding these proteins can be used as probes for the identification
 CC of additional members of this superfamily of receptor molecules. The
 CC proteins may further comprise a second hydrophobic domain at the amino
 CC terminal which comprises a secretion signal sequence which promotes
 CC the intracellular transport of the initially expressed receptor
 CC protein across the golgi membrane. These receptor proteins can be
 CC used to develop agents for the diagnosis and/or treatment of eg.
 CC activin-dependent tumours, for enhancing the survival of brain
 CC neurons, for inducing abortion or twinning in livestock, for
 CC stimulating wound healing, for suppression of growth of TGF-beta
 CC sensitive tumours, for suppressing immune response, for promoting
 CC liver regeneration and for stimulating some immune responses.
 SQ Sequence 513 AA;
 SQ 38 A; 22 R; 20 N; 28 D; 0 B; 19 C; 17 Q; 38 E; 0 Z; 32 G; 15 H;
 SQ 29 I; 46 L; 31 K; 16 M; 20 F; 30 P; 25 S; 24 T; 11 W; 17 Y; 35 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                || || ||
  INCYDRTDCIEKKDSPEVVFCCCEGNMCNEKFSYFPEMEVTQPTSNPVTPKPPYYNILLYSLVPLMLIAGIV
    90      100      110      120      130 X   140 X   150

  ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
    160      170      180      190

```

6. US-08-249-182-7 (1-10)
 R42635 Human interferon receptor.

ID R42635 standard; Protein; 557 AA.
 AC R42635;
 DT 20-APR-1994 (first entry)
 DE Human interferon receptor.
 KW IFN-R; extracellular domain; monoclonal antibody; viral infection;
 KW cell proliferation; allograft rejection; systemic lupus erythematosus;
 KW psoriasis; multiple sclerosis; Behcet's Disease; aplastic anaemia;

KW immunodeficiency; measles virus; interferon-alpha-beta.
 OS Homo sapiens.
 FM Key Location/Qualifiers
 FT Domain 1..436
 FT /label= extracellular_domain
 FT /note= "soluble, immunogenic form of IFN-R"
 PN EP-563487-A.
 PD 06-OCT-1993.
 PF 31-MAR-1992; 400902.
 PR 31-MAR-1992; EP-400902.
 PA (EUBI-) LAB EURO BIOTECHNOLOGIE SA.
 PI Benoit P, Maguire D, Meyer F, Plavec I, Tovey MG;
 DR WPI; 93-312951/40.
 DR P-PSDB; R42635.
 PT Monoclonal antibody to human interferon type-I receptor - having
 PT neutralising activity against human type I interferon, used for
 PT therapy and diagnosis
 PS Disclosure; Fig 3; 21pp; English.
 CC Monoclonal antibodies produced against soluble forms of the human
 CC interferon alpha-beta receptor based on the full-length human IFN-R
 CC sequence are claimed. The antibodies are useful for treatment and
 CC prophylaxis of disorders involving cell proliferation and/or viral
 CC infection.
 SQ Sequence 557 AA;
 SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
 SQ 43 I; 44 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 42 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPEPTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
      260      270      280
  
```

7. US-08-249-182-7 (1-10)

R28496 Sequence of a soluble form of the interferon (IFN)

ID R28496 standard; Protein; 557 AA.
 AC R28496;
 DT 31-MAR-1993 (first entry)
 DE Sequence of a soluble form of the interferon (IFN) receptor
 DE with a high affinity for IFN-alpha and -beta.
 KW Interferon receptor; alpha-interferon; beta-interferon.
 OS Synthetic.
 PN WD9218626-A.
 PD 29-OCT-1992.
 PF 17-APR-1991; F00318.
 PR 17-APR-1991; WD-F00318.
 PA (EUBI-) LAB EURO BIOTECHNOLOGIE.
 PI Eid P, Gresser I, Lutfalla G, Meyer F, Mogensen KE,
 PI Tovey M, Uze G;
 DR WPI; 92-382110/46.
 DR N-PSDB; Q30533.
 PT Water soluble polypeptide(s) strongly bind interferon(s) alpha
 PT and beta - useful as immunosuppressants, for treating auto-immune
 PT diseases and transplant rejection
 PS Claim 3; Fig 2; 58pp; English.
 CC DNA encoding the water-soluble polypeptide with a high affinity for
 CC IFN-alpha and -beta is isolated by PCR, using appropriate

CC oligonucleotides as primers and cloned cDNA as template. For example,
 CC bacteriophage lambda ZAP, containing the entire coding sequence of
 CC the IFN-alpha and -beta receptor (Q30533), was incubated with oligos
 CC Q30534 and Q30535. R28496 represents the complete receptor. R28495
 CC lacks the transmembrane and cytoplasmic domains. Both forms bind
 CC IFN in the same way as antibodies so are immunosuppressants e.g. for
 CC treating autoimmune diseases and graft rejection. They lack the
 CC toxic side-effects of known immunosuppressants such as steroids.
 SQ Sequence 557 AA;
 SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
 SQ 43 I; 45 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 41 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 || ||||

RIENIYSRHKIYKLSPEPTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
 180 190 200 210 220 230 240 250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
 260 270 280

8. US-08-249-182-7 (1-10)
 R14488 Complete interferon-alpha/beta receptor.

ID R14488 standard; Protein; 557 AA.
 AC R14488;
 DT 16-JAN-1992 (first entry)
 DE Complete interferon-alpha/beta receptor.
 KW IFN; autoimmune disease; graft rejection; histocompatibility.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 437..457
 FT /label= transmembrane
 FT Domain 458..557
 FT /label= cytoplasmic
 PN FR2657881-A.
 PD 09-AUG-1991.
 PF 05-FEB-1990; 001298.
 PR 05-FEB-1990; FR-001298.
 PA (EUBI-) LAB EURO BIOTECHNO.
 PI Eid P, Gresser I, Lutfalla G, Meyer F, Mogensen KE;
 PI Tovey MG, Uze G;
 DR WPI; 91-319778/44.
 DR N-PSDB; Q14240.
 PT New water-soluble polypeptide(s) with affinity for IFN-alpha and
 PT beta - used to treat e.g. lupus erythematosus, Behcet's disease,
 PT aplastic anaemia, diabetes mellitus, rheumatoid arthritis, etc.
 PS Disclosure; Page 47; 52pp; French.
 CC The invention covers derivatives of the interferon-alpha and/or beta
 CC receptor obtained by deleting the transmembrane and cytoplasmic domains
 CC of the native receptor or by substitution. Potentially immunogenic
 CC epitopes are eliminated and the deriv. can be secreted from
 CC transformed cells. Soluble deriv.s block the activity of IFN alpha/beta
 CC and can be used to treat autoimmune diseases or to inhibit graft
 CC rejection. See also Q14239.
 SQ Sequence 557 AA;
 SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
 SQ 43 I; 44 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 42 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
 Residue Identity = 60% Matches = 6 Mismatches = 4

Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPEPTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGQNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
      260      270      280

```

9. US-08-249-182-7 (1-10)

R11958 Human alpha-interferon receptor protein.

ID R11958 standard; Protein; 557 AA.
AC R11958;
DT 18-JUL-1991 (first entry)
DE Human alpha-interferon receptor protein.
KW Human alpha IFN; IFN agonists; antiviral; anti tumour agent;
KW drug targetting.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..27
FT /label= signal peptide
PN W09105862-A.
PD 02-MAY-1991.
PF 19-OCT-1990; F00758.
PR 20-OCT-1989; FR-013770.
PA (CNRS) CNRS CENT NAT RECH SCI.
PI Mogensen KE, Uze G, Lutfalla G, Gresser I;
DR WPI; 91-148740/20.
DR N-PSDB; Q11701.
PT New human alpha-interferon receptor protein - useful for testing
PT interferon agonists and in treatment or diagnosis
PS Disclosure; fig 4; 30pp; French.
CC This recombinant human alpha interferon (IFN) receptor protein is
CC useful for the testing of IFN agonists and for treatment and diag-
CC nosis of viral diseases and tumours. Antibodies raised against
CC this protein can be used for blocking the receptor when required,
CC eg where overexpression of alpha-IFN is harmful. The Abs are
CC also useful for eg drug targetting. Variants of the protein,
CC having residue 164 (Thr) replaced by Arg and an Asp inserted
CC between residues 479 and 480, are also useful.
SQ Sequence 557 AA;
SQ 26 A; 13 R; 35 N; 25 D; 0 B; 11 C; 23 Q; 42 E; 0 Z; 20 G; 9 H;
SQ 43 I; 44 L; 44 K; 6 M; 26 F; 26 P; 51 S; 36 T; 13 W; 22 Y; 42 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.21
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPEPTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGQNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFKRNPGNHLYKWKQIPDCENVKT
      260      270      280

```

10. US-08-249-182-7 (1-10)

R10607 Peptide with motilin-like activity (J).

ID R10607 standard; Protein; 13 AA.

AC R10607;
 DT 18-APR-1991 (first entry)
 DE Peptide with motilin-like activity (J).
 KW Motilin; activity; gastrointestinal disorder; drug.
 OS Synthetic.
 PN J02311495-A.
 PD 27-DEC-1990.
 PF 24-MAY-1989; 128911.
 PR 24-MAY-1989; JP-128911.
 PA (SANW) SANWA KAGAKU KENKYUSHO.
 DR WPI; 91-047299/07.
 PT Polypeptide(s) with motilin-like activity - used as active
 PT component of drug for treating gastrointestinal disorder
 PS Disclosure; Page 4; 7pp; Japanese.
 CC Compared with motilin, the peptide chain is considerably shorter.
 CC Chemical synthesis is easy and cheap. The peptide is used
 CC as active component in a drug for treating gastrointestinal disorders.
 CC See also R10598-R10611.
 SQ Sequence 13 AA;
 SQ 0 A; 1 R; 0 N; 0 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
|| |  ||
FVPIFTYGELKRRK
X      10

```

11. US-08-249-182-7 (1-10)
 R10606 Peptide with motilin-like activity (I).

ID R10606 standard; Protein; 14 AA.
 AC R10606;
 DT 18-APR-1991 (first entry)
 DE Peptide with motilin-like activity (I).
 KW Motilin; activity; gastrointestinal disorder; drug.
 OS Synthetic.
 PN J02311495-A.
 PD 27-DEC-1990.
 PF 24-MAY-1989; 128911.
 PR 24-MAY-1989; JP-128911.
 PA (SANW) SANWA KAGAKU KENKYUSHO.
 DR WPI; 91-047299/07.
 PT Polypeptide(s) with motilin-like activity - used as active
 PT component of drug for treating gastrointestinal disorder
 PS Disclosure; Page 4; 7pp; Japanese.
 CC Compared with motilin, the peptide chain is considerably shorter.
 CC Chemical synthesis is easy and cheap. The peptide is used
 CC as active component in a drug for treating gastrointestinal disorders.
 CC See also R10598-R10611.
 SQ Sequence 14 AA;
 SQ 0 A; 1 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
|| |  ||

```

12. US-08-249-182-7 (1-10)

R10605 Peptide with motilin-like activity (H).

ID R10605 standard; Protein; 15 AA.
AC R10605;
DT 18-APR-1991 (first entry)
DE Peptide with motilin-like activity (H).
KW Motilin; activity; gastrointestinal disorder; drug.
OS Synthetic.
PN J02311495-A.
PD 27-DEC-1990.
PF 24-MAY-1989; 128911.
PR 24-MAY-1989; JP-128911.
PA (SANW) SANWA KAGAKU KENKYUSHO.
DR WPI; 91-047299/07.
PT Polypeptide(s) with motilin-like activity - used as active
PT component of drug for treating gastrointestinal disorder
PS Disclosure; Page 4; 7pp; Japanese.
CC Compared with motilin, the peptide chain is considerably shorter.
CC Chemical synthesis is easy and cheap. The peptide is used
CC as active component in a drug for treating gastrointestinal disorders.
CC See also R10598-R10611.
SQ Sequence 15 AA;
SQ 0 A; 2 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 1 E; 0 Z; 1 G; 0 H;
SQ 1 I; 1 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
Residue Identity = 50% Matches = 5 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X X
VPPFENIELY
|| | ||
FVPIFTYGELEGRKRK
X 10

13. US-08-249-182-7 (1-10)

R10601 Peptide with motilin-like activity (D).

ID R10601 standard; Protein; 15 AA.
AC R10601;
DT 18-APR-1991 (first entry)
DE Peptide with motilin-like activity (D).
KW Motilin; activity; gastrointestinal disorder; drug.
OS Synthetic.
PN J02311495-A.
PD 27-DEC-1990.
PF 24-MAY-1989; 128911.
PR 24-MAY-1989; JP-128911.
PA (SANW) SANWA KAGAKU KENKYUSHO.
DR WPI; 91-047299/07.
PT Polypeptide(s) with motilin-like activity - used as active
PT component of drug for treating gastrointestinal disorder
PS Disclosure; Page 4; 7pp; Japanese.
CC Compared with motilin, the peptide chain is considerably shorter.
CC Chemical synthesis is easy and cheap. The peptide is used
CC as active component in a drug for treating gastrointestinal disorders.
CC See also R10598-R10611.
SQ Sequence 15 AA;
SQ 0 A; 1 R; 0 N; 0 D; 0 B; 0 C; 2 Q; 2 E; 0 Z; 1 G; 0 H;
SQ 1 I; 2 L; 0 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
|| |  ||
FVPIFTYGELQRLQE
X      10

```

14. US-08-249-182-7 (1-10)

R10604 Peptide with motilin-like activity (G).

ID R10604 standard; Protein; 17 AA.
 AC R10604;
 DT 18-APR-1991 (first entry)
 DE Peptide with motilin-like activity (G).
 KW Motilin; activity; gastrointestinal disorder; drug.
 OS Synthetic.
 PN J02311495-A.
 PD 27-DEC-1990.
 PF 24-MAY-1989; 128911.
 PR 24-MAY-1989; JP-128911.
 PA (SANW) SANWA KAGAKU KENKYUSHO.
 DR WPI; 91-047299/07.
 PT Polypeptide(s) with motilin-like activity - used as active
 PT component of drug for treating gastrointestinal disorder
 PS Disclosure; Page 4; 7pp; Japanese.
 CC Compared with motilin, the peptide chain is considerably shorter.
 CC Chemical synthesis is easy and cheap. The peptide is used
 CC as active component in a drug for treating gastrointestinal disorders.
 CC See also R10598-R10611.
 SQ Sequence 17 AA;
 SQ 0 A; 2 R; 0 N; 0 D; 0 B; 0 C; 2 Q; 1 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 2 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
VPPFENIELY
|| |  ||
FVPIFTYGELQRLQKRK
X      10

```

15. US-08-249-182-7 (1-10)

R10603 Peptide with motilin-like activity (F).

ID R10603 standard; Protein; 18 AA.
 AC R10603;
 DT 18-APR-1991 (first entry)
 DE Peptide with motilin-like activity (F).
 KW Motilin; activity; gastrointestinal disorder; drug.
 OS Synthetic.
 PN J02311495-A.
 PD 27-DEC-1990.
 PF 24-MAY-1989; 128911.
 PR 24-MAY-1989; JP-128911.
 PA (SANW) SANWA KAGAKU KENKYUSHO.
 DR WPI; 91-047299/07.
 PT Polypeptide(s) with motilin-like activity - used as active
 PT component of drug for treating gastrointestinal disorder

PS Disclosure, Page 477pp, Japanese.
 CC Compared with motilin, the peptide chain is considerably shorter.
 CC Chemical synthesis is easy and cheap. The peptide is used
 CC as active component in a drug for treating gastrointestinal disorders.
 CC See also R10598-R10611.
 SQ Sequence 18 AA;
 SQ 0 A; 2 R; 0 N; 0 D; 0 B; 0 C; 2 Q; 2 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 2 L; 2 K; 0 M; 2 F; 1 P; 0 S; 1 T; 0 W; 1 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.37
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

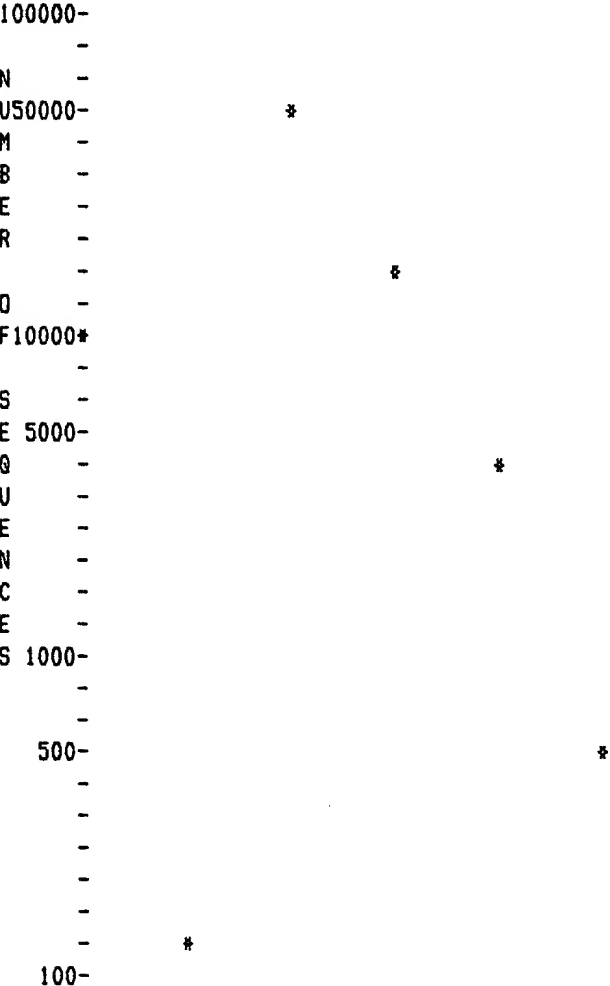
X X
 VPPFENIELY
 || ||
 FVPIFTYGELQRLQEKRK
 X 10
 > 0 <
 0| 0 IntelliGenetics
 > 0 <

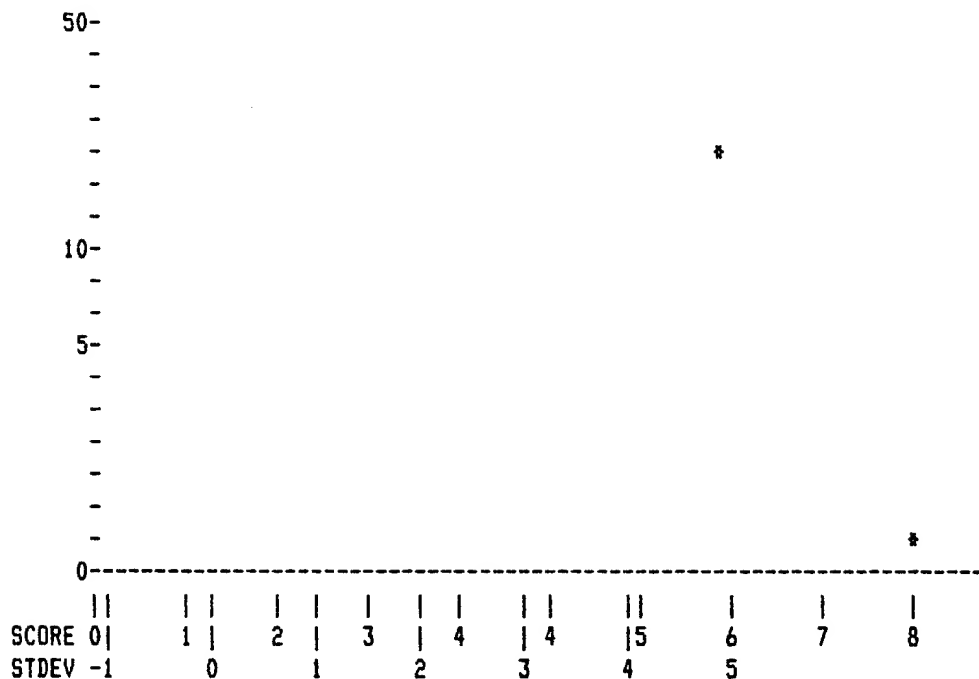
FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_7p.res made by on Thu 22 Sep 94 10:44:49-PDT.

Query sequence being compared:US-08-249-182-7 (1-10)
 Number of sequences searched: 70848
 Number of scores above cutoff: 3987

Results of the initial comparison of US-08-249-182-7 (1-10) with:
 Data bank : PIR 41, all entries





PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.98

Times:	CPU	Total Elapsed
	00:01:25.99	00:01:36.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	3987

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

*** 6 standard deviations above mean ***

##note sequence extracted from NCBI backbone
 SUMMARY #length 114 #checksum 7335
 SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 6.11
 Residue Identity = 80% Matches = 8 Mismatches = 2
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                ||||| ||
DITLVPGTLGRDIEHLTSLDFFRVNSMGTVFVGYGPTFKGGQPLWITATKSPPFENINLYYDVPWNETIPEE
      20      30      40      50      60 X      70 X      80

VTXPNYLQAEVSYPAFKPXL DVYKWHVAAN
      90      100      110

```

2. US-08-249-182-7 (1-10)

R3HS12 ribosomal protein HS12 - Haloarcula marismortui

ENTRY R3HS12 #type complete
 TITLE ribosomal protein HS12 - Haloarcula marismortui
 ALTERNATE_NAMES ribosomal protein E1.3
 ORGANISM #formal_name Haloarcula marismortui
 DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
 30-Jun-1993
 ACCESSIONS S00183; C24304
 REFERENCE S00182
 #authors Kimura, J.; Arndt, E.; Kimura, M.
 #journal FEBS Lett. (1987) 224:65-70
 #title Primary structures of three highly acidic ribosomal proteins
 S6, S12 and S15 from the archaeobacterium Halobacterium
 marismortui.
 #cross-references MUID:88055606
 #accession S00183
 ##molecule_type protein
 ##residues 1-147 ##label KIM
 ##note the source is designated as Halobacterium marismortui
 REFERENCE A24304
 #authors Shoham, M.; Dijk, J.; Reinhardt, R.; Wittmann-Liebold, B.
 #journal FEBS Lett. (1986) 204:323-330
 #title Purification and characterization of ribosomal proteins from
 the 30 S subunit of the extreme halophile Halobacterium
 marismortui.
 #accession C24304
 ##molecule_type protein
 ##residues 1-11,'E',13-14,'I',16-18,'I',20-21 ##label SHO
 ##note the source is designated as Halobacterium marismortui
 CLASSIFICATION #superfamily rat ribosomal protein S19
 KEYWORDS protein biosynthesis; ribosome
 SUMMARY #length 147 #molecular-weight 16438 #checksum 4777
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

X      10
VPPFENIELY
||| |||
ATLYDVPPEELIEALTETLADEDWAEFTKTGVDRLEPPEQEDFWTRRAASLLRKVAVDGPVGVNA
X 10 X 20 30 40 50 60

```

1. A42329	autotaxin - human (fragments)	114	8	8.11	0
**** 4 standard deviations above mean ****					
2. R3HS12	ribosomal protein HS12 - Halo	147	6	6	4.07 0
3. S31985	rasC protein - slime mold (Di	189	6	6	4.07 0
4. S15960	hypothetical protein 1 - yeas	233	6	6	4.07 0
5. J01743	hypothetical 33.6K protein -	287	6	6	4.07 0
6. S33518	probable nucleotide-binding p	369	6	6	4.07 0
7. S37563	actin-related protein - yeast	420	6	6	4.07 0
8. A49193	type II activin receptor ActR	513	6	6	4.07 0
9. S27258	activin receptor type II - ra	513	6	6	4.07 0
10. S18908	activin receptor precursor -	513	6	6	4.07 0
11. S22345	activin receptor - human	513	6	6	4.07 0
12. A39896	activin receptor precursor -	513	6	6	4.07 0
13. J01486	activin receptor II precursor	513	6	6	4.07 0
14. S17112	interferon alpha/beta recepto	545	6	6	4.07 0
15. A32694	interferon alpha receptor pre	557	6	6	4.07 0
16. E35216	FPD5 protein - fowlpox virus	791	6	6	4.07 0
17. S16133	dimethylglycine dehydrogenase	857	6	6	4.07 0
18. S37310	protoporphyrin IX magnesium c	1379	6	6	4.07 0
19. J01917	polyprotein - parsnip yellow	3027	6	6	4.07 0
**** 3 standard deviations above mean ****					
20. B36558	H+/K+-exchanging ATPase (EC 3	17	5	5	3.05 0
21. A24228	H+/K+-exchanging ATPase (EC 3	17	5	5	3.05 0
22. A60313	motilin - dog	22	5	5	3.05 0
23. S00189	motilin - dog	22	5	5	3.05 0
24. MSPG	motilin - pig	22	5	5	3.05 0
25. S30335	alcohol/aldehyde oxidoreducta	41	5	5	3.05 0
26. S01587	hypothetical protein 50 - liv	50	5	5	3.05 0
27. A05031	hypothetical protein 50 - liv	50	5	5	3.05 0
28. A27210	hypothetical protein (CDII 5'	59	5	5	3.05 0
29. WMVZK8	K8 protein - vaccinia virus (64	5	5	3.05 0
30. J01197	hypothetical 7.6K protein (re	69	5	5	3.05 0
31. N2EP2D	long neurotoxin 2 - black mam	72	5	5	3.05 0
32. N2EP1D	long neurotoxin 1 - black mam	72	5	5	3.05 0
33. S31018	gene 73 protein - Mycobacteri	85	5	5	3.05 0
34. S00740	hypothetical protein - Methan	96	5	5	3.05 0
35. S06415	beta-2-microglobulin - rat	99	5	5	3.05 0
36. MGRBB2	beta-2-microglobulin - rabbit	99	5	5	3.05 0
37. PC1125	c-ros protein - mouse (fragme	100	5	5	3.05 0
38. Z2BPT9	gene 49.2 protein - phage T4	106	5	5	3.05 0
39. A37203	lens fiber membrane major int	112	5	5	3.05 0
40. A30329	promotilin precursor - human	114	5	5	3.05 0

1. US-08-249-182-7 (1-10)

A42329 autotaxin - human (fragments)

```

ENTRY      A42329      #type fragments
TITLE      autotaxin - human (fragments)
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
           08-May-1993
ACCESSIONS A42329
REFERENCE  A42329
#authors   Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
           Cioce, V.; Schiffmann, E.; Liotta, L.A.
#journal   J. Biol. Chem. (1992) 267:2524-2529
#title     Identification, purification, and partial sequence analysis
           of autotaxin, a novel motility-stimulating protein.
#cross-references MUID:92129337
#accession A42329
##status   preliminary
##molecule_type protein
##residues 1-114 ##label STR
##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
           NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;

```

3. US-08-249-182-7 (1-10)

S31985 rasC protein - slime mold (Dictyostelium discoideum)

ENTRY S31985 #type complete
 TITLE rasC protein - slime mold (Dictyostelium discoideum)
 ORGANISM #formal_name Dictyostelium discoideum
 DATE 03-Mar-1994; #sequence_revision 03-Mar-1994; #text_change 03-Mar-1994
 ACCESSIONS S31985
 REFERENCE S31985
 #authors Daniel, J.M.; Bush, J.; Cardelli, J.; Spiegelman, G.B.; Weeks, G.
 #submission submitted to the EMBL Data Library, December 1992
 #accession S31985
 ##status preliminary
 ##residues 1-189 ##label DAN
 ##cross-references EMBL:Z18926
 SUMMARY #length 189 #molecular-weight 21496 #checksum 4479
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 ||| | ||
 GFLIVYSIISRASFEAVTTFREQILRVKDLSTYPIVIIGNKADLPDKDRKVPPECKELAKSFGAPFLETS
 80 90 100 110 120 130 140

KSRVNVVEEAFFTLVREIKRWNPQNEMLPPKRGCI
 150 160 170 180

4. US-08-249-182-7 (1-10)

S15960 hypothetical protein 1 - yeast (Saccharomyces kluyveri)

ENTRY S15960 #type complete
 TITLE hypothetical protein 1 - yeast (Saccharomyces kluyveri) plasmid pSKL
 ORGANISM #formal_name Saccharomyces kluyveri
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change 21-Nov-1993
 ACCESSIONS S15960
 REFERENCE S15960
 #authors Hishinuma, F.; Hirai, K.
 #journal Mol. Gen. Genet. (1991) 226:97-106
 #title Genome organization of the linear plasmid, pSKL, isolated from Saccharomyces kluyveri.
 #cross-references MUID:91238725
 #accession S15960
 ##status preliminary
 ##residues 1-233 ##label HIS
 ##cross-references EMBL:X54850
 SUMMARY #length 233 #molecular-weight 27705 #checksum 5808
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || ||||
 FSYDFEPPEMELFVSASFNNKEVYEDQLSKAKEIGYFYKYLDIDSESSLEVPLKENIETDFGFDPTNTVYGEN
 130 140 150 160 170 X 180 X 190

ELCEEYERKFFRSFEFNKGIEGVKELDEVVRRMGGI
200 210 220 230

5. US-08-249-182-7 (1-10)

J01743 hypothetical 33.6K protein - rabbit fibrona virus

ENTRY J01743 #type complete
TITLE hypothetical 33.6K protein - rabbit fibrona virus
ALTERNATE_NAMES H2R protein; protein kinase homolog
ORGANISM #formal_name rabbit fibrona virus, Shope fibrona virus
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
30-Sep-1993
ACCESSIONS J01743
REFERENCE J01741
#authors Massung, R.F.; McFadden, G.; Moyer, R.W.
#journal J. Gen. Virol. (1992) 73:2903-2911
#title Nucleotide sequence analysis of a unique near-terminal region
of the tumorigenic poxvirus, Shope fibrona virus.
#accession J01743
##molecule_type DNA
##residues 1-287 ##label MAS
SUMMARY #length 287 #molecular-weight 33561 #checksum 2007
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
VPPFENIELY
|||| ||
LDMALDCCTGLVNVYKYTKPKPYEHLSSVNFVDEYTTKILCHGLEKTLSPVPFKNINMLAYRPRILLDIF
120 130 140 150 160 170 180 190
SKYTEKDDVYSLGVVLWEIFTGCVFPDHATSTDIYDAV
200 210 220

6. US-08-249-182-7 (1-10)

S33518 probable nucleotide-binding protein - Acholeplasma

ENTRY S33518 #type complete
TITLE probable nucleotide-binding protein - Acholeplasma laidlawii
ORGANISM #formal_name Acholeplasma laidlawii
DATE 03-Mar-1994; #sequence_revision 03-Mar-1994; #text_change
03-Mar-1994
ACCESSIONS S33518
REFERENCE S33518
#authors Boyer, M.J.; Jarhede, T.K.; Tegnan, V.; Wieslander, A.
#submission submitted to the EMBL Data Library, June 1993
#accession S33518
##status preliminary
##residues 1-369 ##label BOY
##cross-references EMBL:Z22875
SUMMARY #length 369 #molecular-weight 42141 #checksum 5061
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X X
VPPFENIELY
|||| ||

MMILVTGAKGPIGANTIAELKNGAFENITLYDIDSTIESELDLYCWSADFIPLAGVNRPRDEREFELGNFGP
10 20 X 30 40 50 60 70

TSTLLDKLK
80

7. US-08-249-182-7 (1-10)

S37563 actin-related protein - yeast (*Saccharomyces cerev*

ENTRY S37563 #type complete
TITLE actin-related protein - yeast (*Saccharomyces cerevisiae*)
ORGANISM #formal_name *Saccharomyces cerevisiae*
DATE 08-Dec-1993; #sequence_revision 08-Dec-1993; #text_change
08-Dec-1993
ACCESSIONS S37563
REFERENCE S37563
#authors Harata, M.; Karwan, A.; Wintersberger, U.
#submission submitted to the EMBL Data Library, September 1993
#accession S37563
##status preliminary
##residues 1-420 ##label HAR
##cross-references EMBL:X75317
SUMMARY #length 420 #molecular-weight 47649 #checksum 8418
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
VPPFENIELY
|| || ||
RTKPSGVNKSDDKVTPTTEEKEQEA VSKSTSPAANSADTPNETGKRPLEEEKPPKENNELIGLADLVSSIMS
260 270 280 290 300 X 310 X 320

SDVDLRATLAHNVVLTGGTSSIPGLSDRLMTELNKILP
330 340 350 360

8. US-08-249-182-7 (1-10)

A49193 type II activin receptor ActRII - rat (fragment)

ENTRY A49193 #type fragment
TITLE type II activin receptor ActRII - rat (fragment)
ORGANISM #formal_name *Rattus norvegicus* #common_name Norway rat
DATE 19-Dec-1993; #sequence_revision 19-Dec-1993; #text_change
29-Jan-1994
ACCESSIONS A49193
REFERENCE A49193
#authors Feng, Z.M.; Madigan, M.B.; Chen, C.L.
#journal Endocrinology (1993) 132:2593-2600
#title Expression of type II activin receptor genes in the male and
female reproductive tissues of the rat.
#cross-references MUID:93279247
#accession A49193
##status preliminary
##molecule_type nucleic acid
##residues 1-513 ##label FEN
##cross-references NCBIN:133008; NCBIP:133009
##note sequence extracted from NCBI backbone
SUMMARY #length 513 #checksum 89
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4

Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                || || ||
    INCYDRTDCIEKKDSPEVYFCCEGNCNEKFSYFPFEMEVTQPTSNPVTPKPPYYNILLSVPLMLIAGIV
      90      100      110      120      130 X      140 X      150

    ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
      160      170      180      190

```

9. US-08-249-182-7 (1-10)

S27258 activin receptor type II - rat

ENTRY S27258 #type complete
 TITLE activin receptor type II - rat
 ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993
 ACCESSIONS S27258
 REFERENCE S27258
 #authors Shinozaki, H.; Ito, I.; Hasegawa, Y.; Nakanura, K.; Igarashi, S.; Nakanura, M.; Miyamoto, K.; Eto, Y.; Ibuki, Y.; Minegishi, T.
 #journal FEBS Lett. (1992) 312:53-56
 #title Cloning and sequencing of a rat type II activin receptor.
 #accession S27258
 ##status preliminary
 ##residues 1-513 ##label SHI
 SUMMARY #length 513 #molecular-weight 57903 #checksum 9412
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                || || ||
    INCYDRTDCIEKKDSPEVYFCCEGNCNEKFSYFPFEMEVTQPTSNPVTPKPPYYNILLSVPLMLIAGIV
      90      100      110      120      130 X      140 X      150

    ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
      160      170      180      190

```

10. US-08-249-182-7 (1-10)

S18908 activin receptor precursor - human

ENTRY S18908 #type complete
 TITLE activin receptor precursor - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change 22-Nov-1993
 ACCESSIONS S18908
 REFERENCE S18908
 #authors Geiser, A.G.
 #submission submitted to the EMBL Data Library, December 1991
 #accession S18908
 ##status preliminary
 ##residues 1-513 ##label GEI
 ##cross-references EMBL:X62381
 SUMMARY #length 513 #molecular-weight 57847 #checksum 205
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10
                                     VPPFENIELY
                                     || || ||
INCYDRTDCVEKKDSPEVYFCCCEGNCNEKFSYFPMEVETQPTSNPVTPKPPYYNILLSLVPLMLIAGIV
   90      100      110      120      130 X      140 X      150

ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
   160      170      180      190
```

11. US-08-249-182-7 (1-10)

S22345 activin receptor - human

ENTRY S22345 #type complete
TITLE activin receptor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Nov-1993; #sequence_revision 22-Nov-1993; #text_change
22-Nov-1993
ACCESSIONS S22345
REFERENCE S22345
#authors Matzuk, M.M.; Bradley, A.
#journal Biochim. Biophys. Acta (1992) 1130:105-108
#title Cloning of the human activin receptor cDNA reveals high
evolutionary conservation.
#cross-references MUID:92182002
#accession S22345
##status preliminary
##residues 1-513 ##label MAT
##cross-references EMBL:X63128
SUMMARY #length 513 #molecular-weight 57847 #checksum 205
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                     X      10
                                     VPPFENIELY
                                     || || ||
INCYDRTDCVEKKDSPEVYFCCCEGNCNEKFSYFPMEVETQPTSNPVTPKPPYYNILLSLVPLMLIAGIV
   90      100      110      120      130 X      140 X      150

ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
   160      170      180      190
```

12. US-08-249-182-7 (1-10)

A39896 activin receptor precursor - mouse

ENTRY A39896 #type complete
TITLE activin receptor precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change
18-Jun-1993
ACCESSIONS A39896
REFERENCE A39896
#authors Mathews, L.S.; Vale, W.W.
#journal Cell (1991) 65:973-982
#title Expression cloning of an activin receptor, a predicted
transmembrane serine kinase.
#cross-references MUID:91256317
#accession A39896


```

##status      preliminary
##molecule_type mRNA
##residues    1-513 ##label MAT
##cross-references GB:M65287
SUMMARY      #length 513 #molecular-weight 57889 #checksum 23
SEQUENCE

Initial Score      =      6  Optimized Score =      6  Significance =  4.07
Residue Identity =   60%  Matches           =      6  Mismatches  =    4
Gaps              =      0  Conservative Substitutions      =    0

                                X      10
                                VPPFENIELY
                                || || ||
INCYDRTDCIEKKDSPEVYFCCCEGNCNEKFSYFPMEVETQPTSNPVTPKPPYYNILLSVPLMLIAGIV
      90      100      110      120      130 X      140 X      150

ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
      160      170      180      190

```

```

13. US-08-249-182-7 (1-10)
    JQ1486      activin receptor II precursor - human

ENTRY          JQ1486      #type complete
TITLE          activin receptor II precursor - human
ORGANISM       #formal_name Homo sapiens #common_name man
DATE           17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
                09-Jun-1994
ACCESSIONS     JQ1486
REFERENCE      JQ1486
#authors       Donaldson, C.J.; Mathews, L.S.; Vale, W.W.
#journal       Biochem. Biophys. Res. Commun. (1992) 184:310-316
#title         Molecular cloning and binding properties of the human type II
                activin receptor.
#cross-references MUID:92231944
#contents      Testis
#accession     JQ1486
##molecule_type mRNA
##residues     1-513 ##label DON
COMMENT        This protein binds activin A.
CLASSIFICATION #superfamily protein kinase homology
KEYWORDS       glycoprotein; membrane protein
FEATURE
  1-19          #domain signal sequence #status predicted #label SIG\
  20-513        #protein activin receptor II #status predicted #label
                MAT\
  136-161       #domain transmembrane #label TM1\
  199-484       #domain protein kinase homology #label KIN\
  43,66,333     #binding_site carbohydrate (Asn) (covalent) #status
                predicted
SUMMARY        #length 513 #molecular-weight 57847 #checksum 205
SEQUENCE

Initial Score      =      6  Optimized Score =      6  Significance =  4.07
Residue Identity =   60%  Matches           =      6  Mismatches  =    4
Gaps              =      0  Conservative Substitutions      =    0

                                X      10
                                VPPFENIELY
                                || || ||
INCYDRTDCVEKKDSPEVYFCCCEGNCNEKFSYFPMEVETQPTSNPVTPKPPYYNILLSVPLMLIAGIV
      90      100      110      120      130 X      140 X      150

ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
      160      170      180      190

```

14. US-08-249-182-7 (1-10)

S17112 interferon alpha/beta receptor - human

ENTRY S17112 #type complete
 TITLE interferon alpha/beta receptor - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 21-Nov-1993; #sequence_revision 21-Nov-1993; #text_change 21-Nov-1993
 ACCESSIONS S17112
 REFERENCE S17112
 #authors Lutfalla, G.; Gardiner, X.Y.Z.; Proudhon, D.; Vielh, E.; Mogensen, X.Y.Z.; Uze, G.
 #submission submitted to the EMBL Data Library, July 1991
 #description The structuree of the human interferon alpha/beta receptor gene.
 #accession S17112
 ##status preliminary
 ##residues 1-545 ##label LUT
 ##cross-references EMBL:X60459
 SUMMARY #length 545 #molecular-weight 62169 #checksum 672
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || ||||
RIENIYSRHKIYKLSPEPTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFQVQWLHAFLEKRNPGNHLYKWKQIPDCENVKT
      260      270      280

```

15. US-08-249-182-7 (1-10)

A32694 interferon alpha receptor precursor - human

ENTRY A32694 #type complete
 TITLE interferon alpha receptor precursor - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 30-Sep-1993
 ACCESSIONS A32694
 REFERENCE A32694
 #authors Uze, G.; Lutfalla, G.; Gresser, I.
 #journal Cell (1990) 60:225-234
 #title Genetic transfer of a functional human interferon alpha receptor into mouse cells: cloning and expression of its cDNA.
 #cross-references MUID:90124632
 #accession A32694
 ##status preliminary
 ##molecule_type mRNA
 ##residues 1-557 ##label UZE
 ##cross-references GB:J03171
 SUMMARY #length 557 #molecular-weight 63525 #checksum 7035
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.07
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

VPPFENIELY

|| ||||

RIENIYSRHKIYKLSPEITYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVQNGNYVLKWDY
180 190 200 210 220 230 240 250

TYANMTFGVQWLHAFKRNPGNHLKWKQIPDCENVKT
260 270 280

> D <
0| |0 IntelliGenetics
> D <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_7s.res made by on Thu 22 Sep 94 10:35:07-PDT.

Query sequence being compared:US-08-249-182-7 (1-10)
Number of sequences searched: 36000
Number of scores above cutoff: 3802

Results of the initial comparison of US-08-249-182-7 (1-10) with:
Data bank : Swiss-Prot 28, all entries

100000-
-
N -
U50000-
M -
B -
E -
R -
-
D -
F10000-
-
S -
E 5000-
Q -
U -
E *
N -
C -
E -
S 1000-
-
-
500-
-
-
-
-
-
100-
-
-
50-
-
-
-
10-
-

*

*

*

*

*

*

```

-
-
5-
-
-
-
-
-
-
-
0-----
SCORE 0| 1 1 2 3 3 4 5 5 6
STDEV -1 0 1 2 3

```

PARAMETERS

```

Similarity matrix      Unitary      K-tuple      2
Mismatch penalty      1      Joining penalty      20
Gap penalty            1.00      Window size      5
Gap size penalty      0.05
Cutoff score          0
Randomization group    0

```

```

Initial scores to save      40      Alignments to save      15
Optimized scores to save    0      Display context      50

```

SEARCH STATISTICS

```

Scores:                Mean      Median      Standard Deviation
                        2          3          0.88

```

```

Times:                 CPU          Total Elapsed
                        00:00:50.95      00:00:56.00

```

```

Number of residues:      12496420
Number of sequences searched: 36000
Number of scores above cutoff: 3802

```

```

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.
Cut-off raised to 5.

```

```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

```

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
1. RS12_HALMA	30S RIBOSOMAL PROTEIN HS12 (E	147	6	6	4.57	0
2. RASC_DICDI	RAS-LIKE PROTEIN RASC.	189	6	6	4.57	0
3. YK84_CAEEL	HYPOTHETICAL 43.0 KD PROTEIN	378	6	6	4.57	0
4. AVR2_MOUSE	ACTIVIN RECEPTOR TYPE II PREC	513	6	6	4.57	0
5. AVR2_HUMAN	ACTIVIN RECEPTOR TYPE II PREC	513	6	6	4.57	0
6. INAR_HUMAN	INTERFERON-ALPHA RECEPTOR PRE	557	6	6	4.57	0
7. VD05_FQWP1	92.6 KD PROTEIN.	791	6	6	4.57	0
8. POLG_PYFV1	GENOME POLYPROTEIN (CONTAINS:	3027	6	6	4.57	0
**** 3 standard deviations above mean ****						
9. MOTI_CANFA	MOTILIN.	22	5	5	3.42	0

10. ADH1A_HUMAN	ADH1A-DEPENDENT ALCOHOL DEHYDR	41	5	5	3.42	0
11. ATP8_DENBE	ATP SYNTHASE PROTEIN 8 (EC 3.	59	5	5	3.42	0
12. Y7K5_VACCV	HYPOTHETICAL 7.5 KD PROTEIN.	64	5	5	3.42	0
13. NXL3_DENPO	LONG NEUROTOXIN 3 (TOXIN VN2)	72	5	5	3.42	0
14. NXL2_DENPO	LONG NEUROTOXIN 2 (NEUROTOXIN	72	5	5	3.42	0
15. NXL1_DENPO	LONG NEUROTOXIN 1 (NEUROTOXIN	72	5	5	3.42	0
16. VG73_BPML5	GENE 73 PROTEIN (GP73).	85	5	5	3.42	0
17. E310_ADE02	EARLY E3B 10.4 KD PROTEIN PRE	91	5	5	3.42	0
18. YN11_METIV	HYPOTHETICAL PROTEIN IN NIFH2	96	5	5	3.42	0
19. B2MG_RABIT	BETA-2-MICROGLOBULIN.	99	5	5	3.42	0
20. Y492_BPT4	HYPOTHETICAL 12.6 KD PROTEIN	106	5	5	3.42	0
21. MIP_CHICK	LENS FIBER MAJOR INTRINSIC PR	112	5	5	3.42	0
22. MOTI_HUMAN	MOTILIN PRECURSOR.	115	5	5	3.42	0
23. B2MG_CYPCA	BETA-2-MICROGLOBULIN PRECURSO	116	5	5	3.42	0
24. B2MG_BRARE	BETA-2-MICROGLOBULIN.	116	5	5	3.42	0
25. MOTI_PIG	MOTILIN PRECURSOR.	119	5	5	3.42	0
26. B2MG_RAT	BETA-2-MICROGLOBULIN PRECURSO	119	5	5	3.42	0
27. IATR_PIG	INTER-ALPHA-TRYPSIN INHIBITOR	123	5	5	3.42	0
28. MOTI_RABIT	MOTILIN PRECURSOR.	133	5	5	3.42	0
29. YCP4_SYNPY	HYPOTHETICAL 16.1 KD PROTEIN	140	5	5	3.42	0
30. YBEA_ECOLI	HYPOTHETICAL 17.3 KD PROTEIN	155	5	5	3.42	0
31. PETD_SCEDB	CYTOCHROME B6-F COMPLEX SUBUN	160	5	5	3.42	0
32. PETD_PROHO	CYTOCHROME B6-F COMPLEX SUBUN	160	5	5	3.42	0
33. HPPK_BACSU	2-AMINO-4-HYDROXY-6-HYDROXYME	162	5	5	3.42	0
34. DESS_MYXXA	DEVELOPMENT-SPECIFIC PROTEIN	173	5	5	3.42	0
35. DEST_MYXXA	DEVELOPMENT-SPECIFIC PROTEIN	175	5	5	3.42	0
36. YOP0_YEREN	YOP0 PROTEIN PRECURSOR.	182	5	5	3.42	0
37. NRFG_ECOLI	NRFG PROTEIN.	198	5	5	3.42	0
38. CAT4_AGRTU	CHLORAMPHENICOL ACETYLTRANSFE	209	5	5	3.42	0
39. CYB6_WHEAT	CYTOCHROME B6 (EC 1.10.99.1).	215	5	5	3.42	0
40. CYB6_TOBAC	CYTOCHROME B6 (EC 1.10.99.1).	215	5	5	3.42	0

1. US-08-249-182-7 (1-10)

RS12_HALMA 30S RIBOSOMAL PROTEIN HS12 (E1.3).

ID RS12_HALMA STANDARD; PRT; 147 AA.
AC P19952;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE 30S RIBOSOMAL PROTEIN HS12 (E1.3).
OS HALOARCUA MARISMORTUI (HALOBACTERIUM MARISMORTUI).
OC PROKARYOTA; MENDOSICUTES; ARCHAEBACTERIA; HALOBACTERIALES;
OC HALOBACTERIACEAE.
RN [1]
RP SEQUENCE.
RM 88055606
RA KIMURA J., ARNDT E., KIMURA M.;
RL FEBS LETT. 224:65-70(1987).
RN [2]
RP SEQUENCE OF 1-21.
RA SHOHAM M., DIJK J., REINHARDT R., WITTMANN-LIEBOLD B.;
RL FEBS LETT. 204:323-330(1986).
CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
DR PIR; S00183; R3HS12.
DR PROSITE; PS00628; RIBOSOMAL_S19E.
KW RIBOSOMAL PROTEIN.
FT CONFLICT 12 12 I -> E (IN REF. 2).
FT CONFLICT 15 15 L -> I (IN REF. 2).
FT CONFLICT 19 19 L -> I (IN REF. 2).
SQ SEQUENCE 147 AA; 16438 MW; 101142 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 ||| ||
 ATLYDVPPEELIEALTETLADEWDAEFTKTGVDRELPPGEQDFWTRRAASLLRKVAVDGPVGVNA
 X 10 X 20 30 40 50 60

2. US-08-249-182-7 (1-10)

RASC_DICDI RAS-LIKE PROTEIN RASC.

ID RASC_DICDI STANDARD; PRT; 189 AA.
 AC P32253;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE RAS-LIKE PROTEIN RASC.
 GN RASC.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;
 OC EUMYCETOZOA; DICTYOSTELIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA DANIEL J.M., BUSH J., CARDELLI J., SPIEGELMAN G.B., WEEKS G.;
 RL ONCOGENE 9:501-508(1994).
 CC -!- FUNCTION: RAS PROTEINS BIND GDP/GTP AND POSSESS INTRINSIC GTPASE
 CC ACTIVITY.
 DR EMBL; Z18926; DDRASCA.
 DR PIR; S31985; S31985.
 DR DICTYDB; DD05036; RASC.
 KW GTP-BINDING; PRENYLATION; LIPOPROTEIN.
 FT NP_BIND 11 18 GTP (BY SIMILARITY).
 FT NP_BIND 58 62 GTP (BY SIMILARITY).
 FT NP_BIND 117 120 GTP (BY SIMILARITY).
 FT DOMAIN 33 41 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 186 186 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 189 AA; 21496 MW; 181053 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 ||| ||
 GFLIVYSIISRASFEAVTTFREQILRVKDLSTYPIVIIGNKADLPDKRVPPMEGKELAKSFGAPFLETS
 80 90 100 110 120 130 140
 KSRVNVEEAFFTLVREIKRWNPQNEMLPPKKRGCI
 150 160 170 180

3. US-08-249-182-7 (1-10)

YK84_CAEL HYPOTHETICAL 43.0 KD PROTEIN C30A5.4 IN CHROMOSOME

ID YK84_CAEL STANDARD; PRT; 378 AA.
 AC P34350;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 43.0 KD PROTEIN C30A5.4 IN CHROMOSOME III.
 GN C30A5.4.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; ACCELEMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;
 RA ANDERSON K.;
 RL SUBMITTED (FEB-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; L10990; CEC30A5.
 DR WORMPEP; C30A5.4; CE00095.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 378 AA; 43031 MW; 750062 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 II IIII
 AHQMCDFGYWVLSGRKLITIFSAPMYCNFYKNTGCVLKVDETLGIQTVAFVPESENIEKIIIEEMNRVNDISI
 310 320 330 340 350 X 360 X 370

DCIE

4. US-08-249-182-7 (1-10)

AVR2_MOUSE ACTIVIN RECEPTOR TYPE II PRECURSOR (ACTR-II) (EC 2

ID AVR2_MOUSE STANDARD; PRT; 513 AA.
 AC P27038;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE ACTIVIN RECEPTOR TYPE II PRECURSOR (ACTR-II) (EC 2.7.1.-).
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91256317
 RA MATHEWS L.S., VALE W.W.;
 RL CELL 65:973-982(1991).
 CC -!- FUNCTION: RECEPTOR FOR ACTIVIN A, ACTIVIN B, AND INHIBIN A.
 CC INVOLVED IN TRANSMEMBRANE SIGNALING.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: BRAIN, TESTIS, INTESTINE, LIVER, AND KIDNEY.
 CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
 CC PROTEIN KINASES.
 DR EMBL; M65287; MACTR.
 DR PIR; A39896; A39896.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST.
 KW RECEPTOR; TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
 KW TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 513 ACTIVIN RECEPTOR TYPE II.
 FT DOMAIN 20 135 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 136 161 POTENTIAL.
 FT DOMAIN 162 513 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 198 484 CATALYTIC.
 FT NP_BIND 198 206 ATP (BY SIMILARITY).
 FT BINDING 219 219 ATP (BY SIMILARITY).
 FT ACT_SITE 322 322 BY SIMILARITY.
 FT CARBOHYD 43 43 POTENTIAL.
 FT CARBOHYD 66 66 POTENTIAL.
 SQ SEQUENCE 513 AA; 57889 MW; 1345779 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
 Residue Identity = 60% Matches = 6 Mismatches = 4

gaps = 0 Conservative Substitutions = 0
 X 10
 VPPFENIELY
 || || ||
 INCYDRTDCIEKKDSPEVYFCCCEGNNCKEKFYSYFPEMEVTQPTSNPVTPKPPYYNILLYSLVPLMLIAGIV
 90 100 110 120 130 X 140 X 150
 ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
 160 170 180 190

5. US-08-249-182-7 (1-10)

AVR2_HUMAN ACTIVIN RECEPTOR TYPE II PRECURSOR (ACTR-II) (EC 2

ID AVR2_HUMAN STANDARD; PRT; 513 AA.
 AC P27037;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE ACTIVIN RECEPTOR TYPE II PRECURSOR (ACTR-II) (EC 2.7.1.-).
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RM 92182002
 RA MATZUK M.M., BRADLEY A.;
 RL BIOCHIM. BIOPHYS. ACTA 1130:105-108(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RM 92231944
 RA DONALDSON C.J., MATHEWS L.S., VALE W.W.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 184:310-316(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY;
 RA GEISER A.G.;
 RL SUBMITTED (DEC-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: RECEPTOR FOR ACTIVIN A, ACTIVIN B, AND INHIBIN A.
 CC INVOLVED IN TRANSMEMBRANE SIGNALING.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
 CC PROTEIN KINASES.
 DR EMBL; X63128; HSACTREC.
 DR EMBL; X62381; HSACTR.
 DR EMBL; M93415; HSACTIIA.
 DR PIR; S18908; S18908.
 DR PIR; J01486; J01486.
 DR PIR; S22345; S22345.
 DR MIM; 102576; TENTH EDITION.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST.
 KW RECEPTOR; TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
 KW TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 513 ACTIVIN RECEPTOR TYPE II.
 FT DOMAIN 20 135 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 136 161 POTENTIAL.
 FT DOMAIN 162 513 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 198 484 CATALYTIC.
 FT NP_BIND 198 206 ATP (BY SIMILARITY).
 FT BINDING 219 219 ATP (BY SIMILARITY).
 FT ACT_SITE 322 322 BY SIMILARITY.

FT CARBOHYD 43 43 POTENTIAL.
 SQ SEQUENCE 513 AA; 57847 MW; 1347089 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || || ||
 INCYDRITDCVEKKDSPEVYFCCCEGNMCNEKFSYFPMEVTDQPTSNPVTPKPPYYNILLSVPLMLIAGIV
 90 100 110 120 130 X 140 X 150
 ICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPLLGLKPL
 160 170 180 190

6. US-08-249-182-7 (1-10)

INAR_HUMAN INTERFERON-ALPHA RECEPTOR PRECURSOR (IFN-ALPHA-REC

ID INAR_HUMAN STANDARD; PRT; 557 AA.
 AC P17181;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE INTERFERON-ALPHA RECEPTOR PRECURSOR (IFN-ALPHA-REC).
 GN IFNAR.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90124632
 RA UZE G., LUTFALLA G., GRESSER I.;
 RL CELL 60:225-234(1990).
 CC -!- FUNCTION: RECEPTOR FOR INTERFERON ALPHA.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: IFN RECEPTORS ARE PRESENT IN ALL TISSUES AND
 CC EVEN ON THE SURFACE OF MOST IFN-RESISTANT CELLS.
 DR EMBL; J03171; HSIFNRA.
 DR PIR; A32694; A32694.
 DR MIM; 107450; TENTH EDITION.
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 557 INTERFERON-ALPHA RECEPTOR.
 FT DOMAIN 28 436 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 437 457 POTENTIAL.
 FT DOMAIN 458 557 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 79 87 BY SIMILARITY.
 FT DISULFID 199 220 BY SIMILARITY.
 FT CARBOHYD 50 50 POTENTIAL.
 FT CARBOHYD 58 58 POTENTIAL.
 FT CARBOHYD 81 81 POTENTIAL.
 FT CARBOHYD 88 88 POTENTIAL.
 FT CARBOHYD 110 110 POTENTIAL.
 FT CARBOHYD 172 172 POTENTIAL.
 FT CARBOHYD 254 254 POTENTIAL.
 FT CARBOHYD 313 313 POTENTIAL.
 FT CARBOHYD 314 314 POTENTIAL.
 FT CARBOHYD 376 376 POTENTIAL.
 FT CARBOHYD 416 416 POTENTIAL.
 FT CARBOHYD 433 433 POTENTIAL.
 SQ SEQUENCE 557 AA; 63525 MW; 1717510 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57

Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                VPPFENIELY
                                || |||
RIENIYSRHKIYKLSPETTYCLKVKAALLTSWKIGVYSPVHCIKTTVENELPPPENIEVSVGNQNYVLKWDY
180      190      200      210      220      230      240      250

TYANMTFGVQWLHAFKRNPGNHLKWKQIPDCENVKT
      260      270      280
```

7. US-08-249-182-7 (1-10)

VD05_FOWP1 92.6 KD PROTEIN.

ID VD05_FOWP1 STANDARD; PRT; 791 AA.
AC P21969;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE 92.6 KD PROTEIN.
GN FPD5.
OS FOWLPOX VIRUS (STRAIN FP-1).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC AVIPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RM 90324937
RA TARTAGLIA J., WINSLOW J., GOEBEL S., JOHNSON G.P., TAYLOR J.,
RA PAOLETTI E.;
RL J. GEN. VIROL. 71:1517-1524(1990).
CC -!- FUNCTION: INVOLVED IN VIRAL REPLICATION, POSSIBLY IN THE
CC ELONGATION OF DNA.
CC -!- SIMILARITY: TO VACCINIA VIRUS D5 PROTEIN.
DR EMBL; X17202; POFPHIND.
DR PIR; E35216; E35216.
KW DNA REPLICATION; DNA-BINDING; ATP-BINDING.
FT NP_BIND 505 512 ATP (POTENTIAL).
SQ SEQUENCE 791 AA; 92655 MW; 3362422 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
Residue Identity = 60% Matches = 6 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                VPPFENIELY
                                ||| |||
DIKDSIFYQGNDAKKFVCTVSTGYKYEEGINVDDITTELM SILDDI@PKTKENFENRELYEQILSSCLMGT
430      440      450      460      470      X 480      X 490

KQCIFFFYGETATGKSTTKLLKSMHNMFLETGQVIL
500      510      520      530
```

8. US-08-249-182-7 (1-10)

POLG_PYFV1 GENOME POLYPROTEIN (CONTAINS: 22.5 KD PROTEIN; 26

ID POLG_PYFV1 STANDARD; PRT; 3027 AA.
AC 005057;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (CONTAINS: 22.5 KD PROTEIN; 26 KD PROTEIN; 31 KD
DE PROTEIN; PROBABLE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48)).
OS PARSNIP YELLOW FLECK VIRUS (ISOLATE P-121) (PYFV).

UC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; SEQUIVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 93107855
 RA TURNBULL-ROSS A.D., REAVY B., MAYO M.A., MURANT A.F.;
 RL J. GEN. VIROL. 73:3203-3211(1992).
 CC -!- SIMILARITY: SOME, TO THE CMPV AND TBRV POLYPROTEINS.
 DR EMBL; D14066; PYFPOLYP.
 KW POLYPROTEIN; ATP-BINDING; COAT PROTEIN; RNA-DIRECTED RNA POLYMERASE.
 FT NP_BIND 1467 1474 ATP (POTENTIAL).
 FT VARIANT 962 962 T -> I.
 FT VARIANT 1373 1373 L -> F.
 SQ SEQUENCE 3027 AA; 336242 MW; 22418420 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.57
 Residue Identity = 60% Matches = 6 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 ||| |||
 LVMETLNGTTKNQEEPSKDVIAILEELGDCGVVEGILEKRRKELLSQFGVMDPPPFDAIELEPGKAQASVCFST
 1660 1670 1680 1690 1700 1710 1720
 DAFGNPLKNPFVELFGKLRDEFERATKQEMPDDILTKF
 1730 1740 1750 1760

9. US-08-249-182-7 (1-10)

MOTI_CANFA MOTILIN.

ID MOTI_CANFA STANDARD; PRT; 22 AA.
 AC P19863;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
 DE MOTILIN.
 OS CANIS FAMILIARIS (DOG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; CARNIVORA.
 RN [1]
 RP SEQUENCE.
 RM 83195948
 RA POITRAS P., REEVE J.R. JR., HUNKAPILLER M.W., HOOD L.E., WALSH J.H.;
 RL REGUL. PEPT. 5:197-208(1983).
 CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF
 CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES
 CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.
 CC -!- RESIDUE 1 MAY ALSO BE LYS OR SER.
 DR PIR; S00189; S00189.
 DR PIR; A60313; A60313.
 KW HORMONE.
 FT UNSURE 1 1
 SQ SEQUENCE 22 AA; 2685 MW; 2092 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 ||| |||
 FVPIFTHSELQKIREKERNKGQ
 X 10 20

10. US-08-249-182-7 (1-10)

ADHN_AMEYMDMA-DEPENDENT ALCOHOL DEHYDROGENASE (EC1.1.99.-)

ID ADHN_AMEYMDMA-DEPENDENT ALCOHOL DEHYDROGENASE (EC1.1.99.-)
 AC P80175;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE MDMA-DEPENDENT ALCOHOL DEHYDROGENASE (EC 1.1.99.-) (MDMA-ADH)
 DE (FRAGMENT).
 OS AMYCOLATOPSIS METHANOLICA.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIB 11946;
 RM 93215662
 RA VAN OPHEM P.W., VAN BEEUMEN J., DUINE J.A.;
 RL EUR. J. BIOCHEM. 212:819-826(1993).
 CC -!- FUNCTION: THIS IS A NOVEL ENZYME, CATALYTICALLY DIFFERENT FROM
 CC OTHER ALCOHOL DEHYDROGENASES. IT IS EFFECTIVE IN OXIDIZING
 CC ETHANOL, OTHER PRIMARY ALCOHOLS AND BENZYLALCOHOL ONLY IN THE
 CC PRESENCE OF P-NITROSO-N,N-DIMETHYLANILINE (NDMA) AS AN ELECTRON
 CC ACCEPTOR. NADH ACTS AS A COFACTOR HERE INSTEAD AS A COENZYME.
 CC -!- CATALYTIC ACTIVITY: ALCOHOL + NDMAH = ALDEHYDE + NDMA(+).
 CC -!- COFACTOR: NAD.
 CC -!- SUBUNIT: HOMOTRIMER.
 DR PIR; S30335; S30335.
 KW OXIDOREDUCTASE; NAD.
 FT UNSURE 41 41
 FT NON_TER 41 41
 SQ SEQUENCE 41 AA; 4589 MW; 9102 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 VPPFENIELY
 ||| ||
 MKTKAAVLHSAGKPFEEIELELDGPREHEVLIKYTATGLXR
 10 X 20 30 40

11. US-08-249-182-7 (1-10)

ATP8_DENBE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE-ASSOC

ID ATP8_DENBE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE-ASSOC
 AC P07513;
 DT 01-APR-1988 (REL. 07, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE-ASSOCIATED PROTEOLIPID
 DE COMPONENT) (A6L).
 OS DENDROTHERA BERTIANA (BERTERO'S EVENING PRIMROSE).
 OG MITOCHONDRION.
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC MYRTALES; ONAGRACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. MUNZIA;
 RA HIESEL R., BRENNICKE A.;
 RL FEBS LETT. 193:164-168(1985).
 RN [2]
 RP ERRATUM.
 RC STRAIN=CV. MUNZIA;
 RA HIESEL R., BRENNICKE A.;

RE FEBS LETT. 201:177-177(1988).
 CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
 DR PIR; A27210; A27210.
 KW HYDROGEN ION TRANSPORT; CF(0) SUBUNIT; MITOCHONDRION.
 SQ SEQUENCE 59 AA; 7009 MW; 18749 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 ||| ||
 MPFLVGLSPFFLYFELIGHFQVEPSPTPTIKGRKWWRLSLFLIFWGERRVKKETKANDC
 10 X 20 30 40 50

12. US-08-249-182-7 (1-10)
 Y7K5_VACCV HYPOTHETICAL 7.5 KD PROTEIN.

ID Y7K5_VACCV STANDARD; PRT; 64 AA.
 AC P18383;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 7.5 KD PROTEIN.
 OS VACCINIA VIRUS (STRAIN WR).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 89067908
 RA BOURSNELL M.E.G., FOULDS I.J., CAMPBELL J.I., BINNS M.M.;
 RL J. GEN. VIROL. 69:2995-3003(1988).
 DR EMBL; D00382; PXVACH3K.
 DR PIR; JS0218; WMVZK8.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 64 AA; 7472 MW; 24024 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || |||
 MDFCKIDVVVSFAHSFDNLINFINTIVPSDIIELHQFLVQSSTTGNIFVKHYNMISPRNIFIY
 10 20 X 30 X 40 50 60

13. US-08-249-182-7 (1-10)
 NLX3_DENPO LONG NEUROTOXIN 3 (TOXIN VN2).

ID NLX3_DENPO STANDARD; PRT; 72 AA.
 AC P25667;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE LONG NEUROTOXIN 3 (TOXIN VN2).
 OS DENDROASPIS POLYLEPIS POLYLEPIS (BLACK MAMBA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.

RA STRYDOM D.J.; HATLEY T.;
 RL S. AFR. J. CHEM. 30:40-48(1977).
 CC -!- LD(50) IS 0.4 MG/KG BY SUBCUTANEOUS INJECTION.
 DR PROSITE; PS00272; SNAKE_TOXIN.
 KW VENOM; NEUROTOXIN; MULTIGENE FAMILY.
 FT DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 14 42 BY SIMILARITY.
 FT DISULFID 27 31 BY SIMILARITY.
 FT DISULFID 46 57 BY SIMILARITY.
 FT DISULFID 58 63 BY SIMILARITY.
 SQ SEQUENCE 72 AA; 7939 MW; 25452 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || |||
 RTCNKTFSDQSKICPPGENICYTKTWCDAWCSRRGKIVELGCAATCPKVKAGVGIKCCSTDNCNLFKFGKPR
 10 X 20 X 30 40 50 60 70

14. US-08-249-182-7 (1-10)

NXL2_DENPO LONG NEUROTOXIN 2 (NEUROTOXIN DELTA) (TOXIN TN2).

ID NXL2_DENPO STANDARD; PRT; 72 AA.
 AC P01397;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE LONG NEUROTOXIN 2 (NEUROTOXIN DELTA) (TOXIN TN2).
 OS DENDROASPIS POLYLEPIS POLYLEPIS (BLACK MAMBA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RA STRYDOM D.J.;
 RL THESIS (1973), UNIVERSITY OF PRETORIA, SOUTH-AFRICA.
 DR PIR; A01668; N2EP2D.
 DR PROSITE; PS00272; SNAKE_TOXIN.
 KW VENOM; NEUROTOXIN; MULTIGENE FAMILY.
 FT DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 14 42 BY SIMILARITY.
 FT DISULFID 27 31 BY SIMILARITY.
 FT DISULFID 46 57 BY SIMILARITY.
 FT DISULFID 58 63 BY SIMILARITY.
 FT CONFLICT 7 7 P -> F (IN PIR DATA BANK).
 SQ SEQUENCE 72 AA; 7948 MW; 25602 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || |||
 RTCNKTPSDQSKICPPGENICYTKTWCDAWCSRRGKIVELGCAATCPKVKAGVEIKCCSTDNCNKFKFGKPR
 10 X 20 X 30 40 50 60 70

15. US-08-249-182-7 (1-10)

NXL1_DENPO LONG NEUROTOXIN 1 (NEUROTOXIN GAMMA) (TOXIN VN1).

ID NXL1_DENPO STANDARD; PRT; 72 AA.
 AC P01396;

DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE LONG NEUROTOXIN 1 (NEUROTOXIN GAMMA) (TOXIN VN1).
 OS DENDROASPIS POLYLEPIS POLYLEPIS (BLACK MAMBA).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; REPTILIA;
 OC LEPIDOSAURIA; SERPENTES.
 RN [1]
 RP SEQUENCE.
 RM 72206049
 RA STRYDOM D.J.;
 RL J. BIOL. CHEM. 247:4029-4042(1972).
 CC -!- LD(50) IS 0.12 MG/KG BY SUBCUTANEOUS INJECTION.
 DR PIR; A01667; N2EPID.
 DR PROSITE; PS00272; SNAKE_TOXIN.
 KW VENOM; NEUROTOXIN; MULTIGENE FAMILY.
 FT DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 14 42 BY SIMILARITY.
 FT DISULFID 27 31 BY SIMILARITY.
 FT DISULFID 46 57 BY SIMILARITY.
 FT DISULFID 58 63 BY SIMILARITY.
 SQ SEQUENCE 72 AA; 8043 MW; 25023 CN;

Initial Score = 5 Optimized Score = 5 Significance = 3.42
 Residue Identity = 50% Matches = 5 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 VPPFENIELY
 || |||
 RTCNKTFSDQSKICPPGENICYTKTHCDAWCSQRGKRVELGCAATCPKVKAGVEIKCCSTDDCDKFQFGKPR
 10 X 20 X 30 40 50 60 70
 > 0 <
 0| |0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

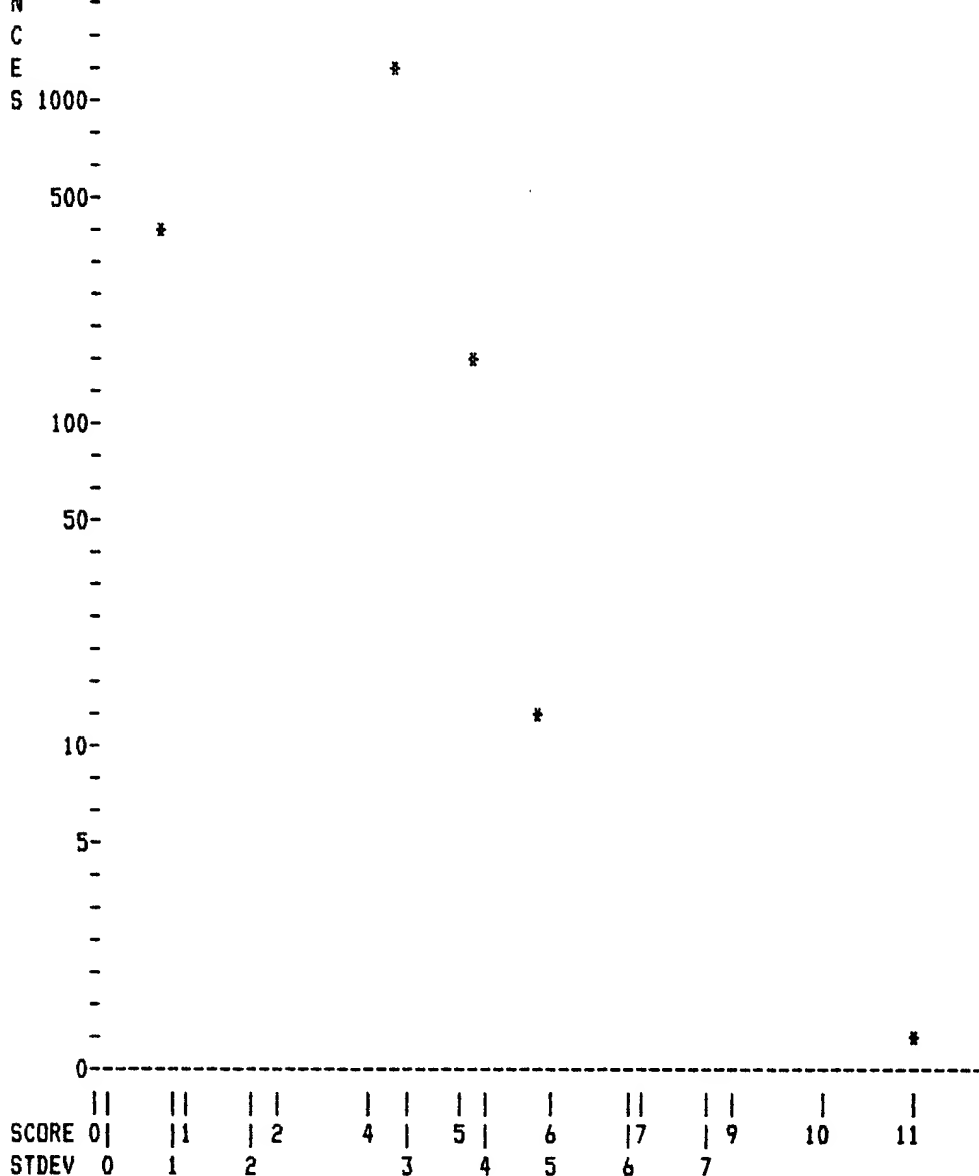
Seq. 8

Results file u249_8a.res made by on Thu 22 Sep 94 10:07:42-PDT.

Query sequence being compared: US-08-249-182-8 (1-11)
 Number of sequences searched: 42145
 Number of scores above cutoff: 4408

Results of the initial comparison of US-08-249-182-8 (1-11) with:
 Data bank : A-GeneSeq 15, all entries

100000-
 -
 N -
 U50000-
 M -
 B -
 E -
 R * *
 -
 O -
 F10000-
 -
 S -
 E 5000- *
 Q -
 U -
 E -



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1	3	1.22
Times:	CPU	Total Elapsed	
	00:00:25.99	00:00:36.00	

Number of residues:	5287517
Number of sequences searched:	42145
Number of scores above cutoff:	4408

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Opt.		Sig.	Frame
			Score	Score		
1. R37450	Autotaxin peptide ATX 100.	11	11	11	8.17	0

The list of other best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 4 standard deviations above mean ****						
2. P70030	Secretory signal sequence of	26	6	6	4.08	0
3. R11208	Retroviral B-epitope containi	31	6	6	4.08	0
4. R36584	Virus neutralising epitope of	33	6	6	4.08	0
5. R14334	HIV-1 amplifier peptide #18.	33	6	6	4.08	0
6. P70033	Secretory signal sequence of	39	6	6	4.08	0
7. P70029	Secretory signal sequence of	50	6	6	4.08	0
8. R13443	FSH inhibiting protein.	264	6	6	4.08	0
9. R05596	Somatomedin carrier protein s	291	6	6	4.08	0
10. P92300	Sequence of human insulin-lik	291	6	6	4.08	0
11. R32501	Beta-adrenergic receptor.	400	6	6	4.08	0
12. R13792	E75B exon B1 polypeptide.	425	6	6	4.08	0
13. R38691	Mitochondria ATPase beta subu	551	6	6	4.08	0
**** 3 standard deviations above mean ****						
14. P98464	Sequence of C. trachomatis se	14	5	5	3.27	0
15. P98452	Sequence of C. trachomatis se	14	5	5	3.27	0
16. P98448	Sequence of C. trachomatis se	14	5	5	3.27	0
17. P98444	Sequence of C. trachomatis se	14	5	5	3.27	0
18. P98468	Sequence of C. trachomatis se	14	5	5	3.27	0
19. P98460	Sequence of C. trachomatis se	14	5	5	3.27	0
20. P82416	Peptide #4 with atrial natriu	20	5	5	3.27	0
21. R12374	Human Factor VII (285-305 + G	23	5	5	3.27	0
22. R34149	CDR2 domain of human V beta 4	24	5	5	3.27	0
23. R24868	Sequence of peptide fragment	28	5	5	3.27	0
24. P91359	Amino acids 482-517 of HIV gl	38	5	5	3.27	0
25. P90504	cDNA from murine cells encodi	69	5	5	3.27	0
26. P94662	Protein sequence for the amin	70	5	6	3.27	0
27. R38924	MIP-1alpha.	74	5	5	3.27	0
28. P91962	Polypeptide encoded by cDNA 5	74	5	5	3.27	0
29. R28371	Papillomavirus E5 protein.	83	5	5	3.27	0
30. R02234	Sequence of the E5 oncogene r	83	5	5	3.27	0
31. P93590	Deduced sequence of MIP-1 alp	92	5	5	3.27	0
32. R12603	SIB 121 intestinal mucin.	95	5	5	3.27	0
33. R13336	HypA protein.	102	5	6	3.27	0
34. R13334	HypA protein.	102	5	6	3.27	0
35. R26953	Human T lymphocyte receptor V	106	5	5	3.27	0
36. R41247	Aphrodisine.	145	5	6	3.27	0
37. R22950	Leech antiplatelet protein.	147	5	5	3.27	0
38. P70494	Sequence of human B-cell grow	159	5	5	3.27	0
39. R42675	Golden hamster Aphrodisin pre	167	5	6	3.27	0
40. R42674	Field hamster Aphrodisin prec	167	5	6	3.27	0

1. US-08-249-182-8 (1-11)

R37450 Autotaxin peptide ATX 100.

ID R37450 standard; peptide; 11 AA.
AC R37450;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 100.

KW Cell motility stimulating; Cancer metastasis; antibody; detection;
 KW immunostains; disease outcome prediction; therapy choice;
 KW cancer therapy; crosslinked toxins.
 OS Synthetic.
 PN US7822043-A.
 PD 01-JAN-1993.
 PF 17-JAN-1992; 822043.
 PR 17-JAN-1992; US-822043.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
 DR WPI; 93-085861/10.
 PT Motility stimulating protein named autotaxin - useful in cancer
 PT diagnosis and therapy
 PS Example; Page 33; 36pp; English.
 CC The sequence is that of autotaxin peptide ATX 100. It may be used to
 CC raise anti-autotaxin antibodies which can be used to diagnose cancer
 CC metastasis and in immunostains of patient samples to detect the
 CC presence of autotaxin. The level of autotaxin in tissue or body
 CC fluids can be used to predict disease outcomes and/or choice of
 CC therapy which may also include autotaxin inhibitors. Autotaxin
 CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
 CC therapy.
 SQ Sequence 11 AA;
 SQ 1 A; 0 R; 0 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 2 G; 0 H;
 SQ 1 I; 1 L; 1 K; 0 M; 0 F; 1 P; 0 S; 2 T; 1 W; 0 Y; 0 V;

Initial Score = 11 Optimized Score = 11 Significance = 8.17
 Residue Identity = 100% Matches = 11 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

      X      X
      GGQPLWITATK
      |||||
      GGQPLWITATK
      X      10
  
```

2. US-08-249-182-8 (1-11)

P70030 Secretory signal sequence of plasmid pSPA12.

ID P70030 standard; protein; 26 AA.
 AC P70030;
 DT 03-FEB-1991 (first entry)
 DE Secretory signal sequence of plasmid pSPA12.
 KW Secretory signal sequence; vector; protein secretion.
 FH Key Location/Qualifiers
 FT Region 17..26
 FT /label=multiple cloning site
 PN EP-244042-A.
 PD 04-NOV-1987.
 PF 29-APR-1987; 200813.
 PR 02-MAY-1986; EP-200774.
 PA (KONN) Gist Brocades NV.
 PI Smith HE, Van Ee JH, Peeters BPH, Bron S, Venema G.
 DR WPI; 87-308298/44.
 PT Plasmid for detecting secretory signal sequence - containing a
 PT multiple cloning site and an open reading frame encoding an enzyme.
 PS Claim 13; Page 28; 38pp; English.
 CC The secretory signal sequence of plasmid pSPA12 comprises a multiple
 CC cloning site with restriction sites in reading frame with the
 CC structural gene. This sequence allows protein secretion, eg in
 CC Bacillus sp., in economically high yields. See also P70028-9 and
 CC P70031-P70041 and N70039-N70041.
 SQ Sequence 26 AA;
 SQ 5 A; 1 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 1 E; 0 Z; 4 G; 0 H;
 SQ 5 I; 3 L; 0 K; 1 M; 0 F; 1 P; 1 S; 1 T; 0 W; 0 Y; 2 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || || ||
MIRGILIAVLGIAIVGGDPLESTAAA
      10      X 20      X

```

3. US-08-249-182-8 (1-11)

R11208 Retroviral B-epitope containing peptide #18 of hyb

ID R11208 standard; Protein; 31 AA.
 AC R11208;
 DT 23-MAY-1991 (first entry)
 DE Retroviral B-epitope containing peptide #18 of hybrid molecule.
 KW retrovirus; env glycoprotein; B-epitope; immunodeficiency virus;
 KW HIV; SIV; AIDS.
 PN W09102544-A.
 PD 07-MAR-1991.
 PF 17-AUG-1990; F00620.
 PR 18-AUG-1989; FR-011044.
 PA (INSP) INST PASTEUR.
 PA (UYCU-) UNIV CURIE P & M PARIS V.
 PI Girard M, Gluckman JC, Bahraoui EM;
 DR WPI; 91-087117/12.
 PT Vaccine compsns. which neutralise human immune deficiency virus -
 PT comprise a B epitope of retro-virus envelope glyco-protein and T
 PT epitope of distinct protein
 PS Claim 6; Page 36; 47pp; French.
 CC The peptide is a specific example of a B-epitope contg. peptide
 CC which can form a hybrid immunogenic molecule with a retroviral T-
 CC epitope. The B-epitope is chosen to be the major neutralisation
 CC epitope of the envelope glycoprotein of a pathogenic retrovirus.
 CC The T-epitope can be derived from a different protein of the
 CC same retrovirus or from the same protein from a different retrovirus.
 CC The hybrid molecule can also contain a minor epitope, especially a
 CC B-epitope from a conserved region of the HIV, SIV, HTLV-1 or
 CC HTLV-II env glycoprotein. The B-epitope-contg. peptide is joined to
 CC the T-epitope using eg tetanus toxin, KLH or HSA.
 CC See also R11191-R11207 and R11209-R11210.
 SQ Sequence 31 AA;
 SQ 2 A; 3 R; 1 N; 0 D; 0 B; 0 C; 3 Q; 0 E; 0 Z; 3 G; 1 H;
 SQ 2 I; 2 L; 2 K; 0 M; 0 F; 2 P; 2 S; 6 T; 0 W; 2 Y; 0 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || | |||
TRPYKNTRQSTPIGLGQALYTTTRTKSIGQAH
      10      X 20      X 30

```

4. US-08-249-182-8 (1-11)

R36584 Virus neutralising epitope of envelope glycoprotei

ID R36584 standard; peptide; 33 AA.
 AC R36584;
 DT 06-SEP-1993 (first entry)
 DE Virus neutralising epitope of envelope glycoprotein of HIV.

KW human immunodeficiency virus; gp120; gp160; EGP; VNE; immunity.
 OS Synthetic.
 PN W09308836-A.
 PD 13-MAY-1993.
 PF 28-OCT-1992; E02459.
 PR 28-OCT-1991; US-782154.
 PR 28-OCT-1991; US-782241.
 PR 28-OCT-1991; US-782252.
 PA (INSP) INST PASTEUR.
 PI Girard M;
 DR WPI; 93-167398/20.
 PT Enhancing immunogenicity of viral envelope glycoprotein - by
 PT co-administration of viral envelope glycoprotein itself, and an
 PT oligopeptide derive.
 PS Disclosure; Page 82; 107pp; English.
 CC A novel method of enhancing the immunogenicity of an envelope
 CC glycoprotein (EGP) of a virus (esp. HIV gp120 or gp160) in a host
 CC comprises admin. to the host at least one EGP of the virus in an amt.
 CC sufficient for priming vaccination and at least one peptide derived
 CC from an amino acid sequence of the EGP (e.g. the sequence shown),
 CC where the peptide comprises at least one virus-neutralisation
 CC epitope (VNE). The complex is able to enhance the induction of
 CC neutralising antibodies to the virus and to confer long lasting
 CC immunity, longer than 6 months.
 CC See also R36567-87.
 SQ Sequence 33 AA;
 SQ 2 A; 3 R; 1 N; 0 D; 0 B; 2 C; 3 Q; 0 E; 0 Z; 3 G; 1 H;
 SQ 2 I; 2 L; 2 K; 0 M; 0 F; 2 P; 2 S; 6 T; 0 W; 2 Y; 0 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || | | ||
CTRPYKNTRQSTPIGLGQALYTTTRKTSIGQAHG
    10      X 20      X 30
  
```

5. US-08-249-182-8 (1-11)

R14334 HIV-1 amplifier peptide #18.

ID R14334 standard; Protein; 33 AA.
 AC R14334;
 DT 03-JAN-1992 (first entry)
 DE HIV-1 amplifier peptide #18.
 KW human immunodeficiency virus; vaccine; human retrovirus; AIDS;
 KW acquired immunodeficiency syndrome; envelope glycoprotein.
 OS Synthetic.
 PN W09114449-A.
 PD 03-OCT-1991.
 PF 16-MAR-1991; E00509.
 PR 19-MAR-1990; US-494749.
 PA (INSP) INST PASTEUR.
 PI Girard M;
 DR WPI; 91-310366/42.
 PT Enhancing immunogenicity of envelope glyco:protein - for use as
 PT vaccine or immuno:therapeutic drug especially against HIV, HTLV-I
 PT and HTLV-II
 PS Claim 12; Page 50; 71pp; English.
 CC This peptide is one example of an HIV-1 amplifier peptide for use in
 CC a composition for enhancing the immunogenicity of an envelope
 CC glycoprotein of a virus. The sequence corresponds to a
 CC neutralisation epitope and enhances the induction of persistent
 CC neutralising antibodies in the host. The amplifier peptide is used

CC in addition to an envelope glycoprotein for priming the induction of
 CC neutralising antibodies. The compositions are particularly
 CC useful for vaccinating against HIV, SIV, HTLV-I and HTLV-II.
 SQ Sequence 33 AA;
 SQ 2 A; 3 R; 1 N; 0 D; 0 B; 2 C; 3 Q; 0 E; 0 Z; 3 G; 1 H;
 SQ 2 I; 2 L; 2 K; 0 M; 0 F; 2 P; 2 S; 6 T; 0 W; 2 Y; 0 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || || ||
  CTRPYKNTRQSTPIGLGGALYTTTRTKSIGQAHK
    10      X 20      X 30
  
```

6. US-08-249-182-8 (1-11)

P70033 Secretory signal sequence of plasmid pSPA31.

ID P70033 standard; protein; 39 AA.
 AC P70033;
 DT 03-FEB-1991 (first entry)
 DE Secretory signal sequence of plasmid pSPA31.
 KW Secretory signal sequence; vector; protein secretion.
 FH Key Location/Qualifiers
 FT Region 30..39
 FT /label=multiple cloning site
 PN EP-244042-A.
 PD 04-NOV-1987.
 PF 29-APR-1987; 200813.
 PR 02-MAY-1986; EP-200774.
 PA (KONN) Gist Brocades NV.
 PI Smith HE, Van Ee JH, Peeters BPH, Bron S, Venema G.
 DR WPI; 87-308298/44.
 PT Plasmid for detecting secretory signal sequence - containing a
 PT multiple cloning site and an open reading frame encoding an enzyme.
 PS Claim 13; Page 28; 38pp; English.
 CC The secretory signal sequence of plasmid pSPA31 comprises a multiple
 CC cloning site with restriction sites in reading frame with the
 CC structural gene. This sequence allows protein secretion, eg in
 CC Bacillus sp., in economically high yields. See also P70028-P70032,
 CC P70034-P70041 and N70039-N70041.
 SQ Sequence 39 AA;
 SQ 5 A; 0 R; 0 N; 2 D; 0 B; 0 C; 0 Q; 8 E; 0 Z; 3 G; 1 H;
 SQ 2 I; 5 L; 0 K; 2 M; 4 F; 2 P; 1 S; 1 T; 0 W; 0 Y; 3 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

      X      10
      GGQPLWITATK
      || || ||
  MDEVHEEEEFEEAPGLFILLFLFVMAVIGGDPLESTAAA
    10      20      30      X
  
```

7. US-08-249-182-8 (1-11)

P70029 Secretory signal sequence of plasmid pSPA3.

ID P70029 standard; protein; 50 AA.
 AC P70029;
 DT 03-FEB-1991 (first entry)
 DE Secretory signal sequence of plasmid pSPA3.

KW Secretory signal sequence; vector; protein secretion.
 FH Key Location/Qualifiers
 FT Region 41..50
 FT /label=multiple cloning site
 PN EP-244042-A.
 PD 04-NOV-1987.
 PF 29-APR-1987; 200813.
 PR 02-MAY-1986; EP-200774.
 PA (KONN) Gist Brocades NV.
 PI Smith HE
 PI Van Ee JH
 PI Peeters BPH
 PI Bron S
 PI Venema G
 DR WPI; 87-308298/44.
 PT Plasmid for detecting secretory signal sequence - containing a
 PT multiple cloning site and an open reading frame encoding an enzyme.
 PS Claim 13; Page 27; 38pp; English.
 CC The secretory signal sequence of plasmid pSPA3 comprises a multiple
 CC cloning site with restriction sites in reading frame with the
 CC structural gene. This sequence allows protein secretion, eg in
 CC Bacillus sp., in economically high yields. (See also P70028,
 CC P70030, P70031, P70032, P70033, P70034, P70035, P70036, P70037,
 CC P70038, P70039, P70040, P70041, N70039, N70040, N70041)
 SQ Sequence 50 AA;
 SQ 6 A; 0 R; 2 N; 2 D; 0 B; 1 C; 0 Q; 4 E; 0 Z; 5 G; 0 H;
 SQ 0 I; 9 L; 4 K; 2 M; 1 F; 2 P; 4 S; 4 T; 0 W; 0 Y; 4 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                X      X
                GGQPLWITATK
                || || ||
MKKMLVLLFSALLLNGCGSGESKANTAETPEVL DVKLTGGDPLESTAAA
      10      20      30      40      50
  
```

8. US-08-249-182-8 (1-11)

R13443 FSH inhibiting protein.

ID R13443 standard; protein; 264 AA.
 AC R13443;
 DT 29-OCT-1991 (first entry)
 DE FSH inhibiting protein.
 KW FSH-IP; endometriosis; contraceptive; porcine follicular fluid;
 KW follicle stimulating hormone.
 OS Sus scrofa domestica.
 PN US5037805-A.
 PD 06-AUG-1991.
 PF 20-MAR-1989; 326151.
 PR 20-MAR-1989; US-326151.
 PA (SALK) SALK INST FOR BIOL STUD.
 PI Ling NC;
 DR WPI; 91-252089/34.
 PT Use of a FSH inhibiting protein - for treating endometriosis and
 PT for use as a male or female contraceptive.
 PS Claim 1; Page 7; 11pp; English.
 CC The protein was isolated and purified from porcine follicular
 CC fluid and may also be prepd. by recombinant DNA techniques. It
 CC can be used for regulating ovulation or fertility in female
 CC mammals, for regulating spermatogenesis in males, and for treating
 CC conditions such as endometriosis which result from an over abun-
 CC dence of FSH or oestrogen. Dosage is pref. 10 ug-1 mg/kg/day.
 SQ Sequence 264 AA;

SD 18 A; 19 R; 6 N; 12 D; 0 B; 18 C; 10 Q; 17 E; 0 Z; 23 G; 6 H;
SQ 6 I; 19 L; 20 K; 2 M; 5 F; 23 P; 27 S; 9 T; 1 W; 9 Y; 14 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                |||  ||
EDTLNKLKFLNVLSPRGVHIPNCDKKGFYKKK@CRPSKGRKRGFCWCVDKYG@PLPGYTTKGKEDVHCYSMQ
      200      210      220      230      240      250      260

```

SK

9. US-08-249-182-8 (1-11)

R05596 Somatomedin carrier protein subunit.

ID R05596 standard; protein; 291 AA.
AC R05596;
DT 31-OCT-1990 (first entry)
DE Somatomedin carrier protein subunit.
KW Human somatomedin carrier protein; LCP2.3; acromegaly;
KW diabetic retinopathy; osteoporosis; ds.
OS Homo sapiens.
PN EP-375438-A.
PD 27-JUN-1990.
PF 21-DEC-1989; 313463.
PR 22-DEC-1988; US-290250.
PA (BIOG-) Biogrowth Inc.
PI Spencer EM, Talkington C;
DR WPI; 90-195533/26.
DR N-PSDB; Q04796.
PT Recombinant DNA molecule -
PT has gene which codes for carrier protein-like polypeptide.
PS Disclosure; p; English.
CC Carrier proteins bind to somatomedin polypeptides, also known as
CC insulin-like polypeptides, they have therapeutic, diagnostic and
CC other applications, such as the inhibition of SM-C in acromegaly,
CC treatment of diabetic retinopathy and wound healing.
SQ Sequence 291 AA;
SD 25 A; 22 R; 6 N; 12 D; 0 B; 18 C; 11 Q; 17 E; 0 Z; 24 G; 7 H;
SQ 6 I; 25 L; 19 K; 3 M; 5 F; 26 P; 27 S; 11 T; 2 W; 9 Y; 16 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                |||  ||
EDTLNHLKFLNVLSPRGVHIPNCDKKGFYKKK@CRPSKGRKRGFCWCVDKYG@PLPGYTTKGKEDVHCYSMQ
      220      230      240      250      260      270      280

```

SK
290

10. US-08-249-182-8 (1-11)

P92300 Sequence of human insulin-like growth factor bindi

ID P92300 standard; protein; 291 AA.
AC P92300;
DT 05-MAR-1990 (first entry)

DE Sequence of human insulin-like growth factor binding protein BP53 from
 KW cDNA clone ibp.118.
 KW Insulin-like growth factor binding protein; BP53; clone ibp.118.
 OS Homo sapiens.
 PN W08909268-A.
 PD 05-OCT-1989.
 PF 10-MAR-1989; U00983.
 PR 22-MAR-1988; US-171623.
 PA (ROYA-) Royal Prince Alfred Hospital (GETH) Genentech Inc.
 PI Baxter RC, Wood WI;
 DR WPI; 89-309533/42.
 DR N-PSDB; N91467.
 PT DNA encoding insulin-like growth factor binding protein - used to
 PT increase insulin-like growth factor circulatory half-life, and as
 PT metabolic regulator.
 PS Disclosure; Fig. 3(1)-3(2); 72pp; English.
 CC BP53 is an acid-stable component of a 125-150 kD glycoprotein complex
 CC contained in human plasma. It carries most of the endogenous IGF's, and
 CC is regulated by growth hormone. To allow more efficient delivery of IGF
 CC to target cells, an AA constituting the transmembrane or membrane-
 CC binding domain of normal IGF-I or -II receptor, or a phospholipid anchor
 CC domain can be introduced at the C-terminal. BP53 can be used to prolong
 CC circulatory half-life of IGF. High levels of BP in adult serum correlate
 CC with high levels of growth hormone.
 SQ Sequence 291 AA;
 SQ 25 A; 22 R; 6 N; 12 D; 0 B; 18 C; 11 Q; 17 E; 0 Z; 25 G; 7 H;
 SQ 6 I; 25 L; 19 K; 3 M; 5 F; 25 P; 27 S; 11 T; 2 W; 9 Y; 16 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

					X	X
					GGGPLWITATK	
EDTLNHLKFLNVLS	SPRGVHIPNC	DKKGFYKKKQ	CRPSKGRKRG	FCWCVDKYG	GPLPGYTT	KGKEDVHCYSMQ
220	230	240	250	260	270	280

SK
 290

11. US-08-249-182-8 (1-11)
 R32501 Beta-adrenergic receptor.

ID R32501 standard; protein; 400 AA.
 AC R32501;
 DT 09-JUN-1993 (first entry)
 DE Beta-adrenergic receptor.
 KW Fat cell specific; BAR; lipolysis; obesity; diagnosis;
 KW thermogenesis; metabolism.
 OS Rattus rattus.
 PN US7783602-A.
 PD 15-DEC-1992.
 PF 11-NOV-1991; 783602.
 PR 01-NOV-1991; US-783602.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Venter CJ,
 DR WPI; 93-067426/08.
 PT Fat cell specific beta- adrenergic receptor polypeptide - used
 PT for diagnosis of obesity due to inactive lipolysis
 PS Disclosure; Page 16; 20pp; English.
 CC A rat intercapsular brown adipose tissue cDNA library was cloned and
 CC probes with DNA probes encoding human beta-1 and rat beta-2
 CC adrenergic receptors under low stringency conditions. Positive
 CC clones were found to be different from the rat and human sequences

CC and contained a single open reading frame of 1200 bp encoding the
CC protein shown, of 400 amino acids and mol. wt. 43 kD. The protein is
CC the fat specific beta-adrenergic receptor and may be used in work on
CC the thermogenesis process. Isolation of the gene for BAR allows the
CC diagnosis and treatment of obesity and the testing of medications
CC for their effectiveness in stimulating the thermogenesis metabolic
CC response in obesity patients.

Sequence 400 AA;
48 A; 32 R; 13 N; 10 D; 0 B; 9 C; 5 Q; 12 E; 0 Z; 28 G; 4 H;
10 I; 50 L; 4 K; 6 M; 18 F; 36 P; 33 S; 26 T; 10 W; 9 Y; 37 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GGQPLWITATK
                                || | |||
MAPWPHKNGSLAFWSDAPTLDPSSAANTSGLPGVPWAAALAGALLALATVGGNLLVITAIARTPRLQITITNVF
      10      20      30      40      50      60      70

VTSLATADLVVGLLVMPGATLALTGHMPLGATGCELW
      80      90     100     110
```

12. US-08-249-182-8 (1-11)
R13792 E75B exon B1 polypeptide.

ID R13792 standard; Protein; 425 AA.
AC R13792;
DT 29-NOV-1991 (first entry)
DE E75B exon B1 polypeptide.
KW Insect steroid receptor; hormone.
OS Drosophila melanogaster.
PN W09113167-A.
PD 05-SEP-1991.
PF 15-FEB-1991; U01189.
PR 26-FEB-1990; US-485749.
PA (STRD) LELAND STANFORD JR UNIV.
PI Hogness DS, Koelle MR, Segraves WA;
DR WPI; 91-281480/38.
DR N-PSDB; Q13573.
PT DNA encoding insect steroid receptors - and ligands, for use as
PT benign inducing factors
PS Disclosure; Page 100; 126pp; English.
CC The amino acid sequence codes for the protein produced by Exon B1
CC which is specific to the E75B transcription unit. Exons 2-5 are
CC shown in E75A (R13791) are also found in E75B but amino acid
CC residues must be increased by 15 to apply to the E75B protein. The
CC E75 proteins show considerable similarity to members of the steroid
CC receptor family. Since the putative hormone binding E domain of the
CC E75 proteins does not show high sequence homology to the known
CC ecdysone receptor (R13793) it is likely that the E75 proteins bind
CC either a terpenoid juvenile hormone or a novel Drosophila hormone.
CC See also R13792-R13794.
Sequence 425 AA;
36 A; 10 R; 25 N; 20 D; 0 B; 13 C; 98 Q; 20 E; 0 Z; 13 G; 22 H;
12 I; 30 L; 18 K; 5 M; 2 F; 17 P; 35 S; 28 T; 1 W; 5 Y; 15 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GGQPLWITATK
                                || | |||
```

10 20 30 40 50 60 70
 HQHQHQHQAKSQQLKQHSALVKLLESAPIKQQQQT
 80 90 100

13. US-08-249-182-8 (1-11)

R38691 Mitochondria ATPase beta subunit.

ID R38691 standard; Protein; 551 AA.
 AC R38691;
 DT 11-NOV-1993 (first entry)
 DE Mitochondria ATPase beta subunit.
 KW Rice; mitochondria; ATPase; beta subunit; male sterility.
 OS Oryza sativa.
 FH Key Location/Qualifiers
 FT Peptide 1..85
 FT /note= "Transit peptide"
 FT Protein 86..551
 FT /note= "ATPase beta subunit"
 PN J05137581-A.
 PD 01-JUN-1993.
 PF 19-NOV-1991; 303251.
 PR 19-NOV-1991; JP-303251.
 PA (MITK) MITSUI TOATSU CHEM INC.
 DR WPI; 93-211307/26.
 DR N-PSDB; 042748.
 PT cDNA of rice mitochondria adenosine triphosphatase beta sub-unit
 PT - for introduction of prod. of foreign gene into mitochondria,
 PT used for recovery of male sterility of rice plant
 PS Claim 1; Page 5-8; 8pp; Japanese.
 CC This sequence is encoded by the rice mitochondria ATPase beta subunit
 CC gene. The gene sequence may be used in the construction of an
 CC artificial mitochondria ATPase beta subunit for introduction into
 CC mitochondria. These may be used for the recovery of male sterile
 CC rice plants. This sequence is given as it is represented in the
 CC specification.
 SQ Sequence 551 AA;
 SQ 57 A; 37 R; 16 N; 28 D; 0 B; 2 C; 23 Q; 34 E; 0 Z; 56 G; 11 H;
 SQ 34 I; 49 L; 21 K; 10 M; 18 F; 30 P; 28 S; 36 T; 0 W; 13 Y; 48 V;

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 GGQPLWITATK
 || ||
 AVHFRDAEGQDVLLFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATDLGGLQERITTTKKSITSVQAIY
 320 330 340 350 360 X 370 X 380
 VPADDLTDPAATTFAHLDATTVLSRQISELGIYPAVDP
 390 400 410 420

14. US-08-249-182-8 (1-11)

P98464 Sequence of C. trachomatis serovar K major outer m

ID P98464 standard; Protein; 14 AA.
 AC P98464;
 DT 06-MAR-1992 (first entry)
 DE Sequence of C. trachomatis serovar K major outer membrane protein (MOMP)
 DE variable domain (VD) K-VDIII encoded by base pairs 742-783
 KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay.

US Chlamydia trachomatis.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97099.
 PT Chlamydia trachomatis genes - used for determn. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 18; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP
 CC specific monoclonal antibodies. The nucleotide, amino acid
 CC sequences and hydrophilicity/charge value analyses will assist in
 CC the selection of appropriate MOMP antigenic determinants to be used
 CC in the construction of synthetic peptides, subunits or recombinant
 CC chlamydial vaccines. This will allow the prodn. of reagents and
 CC methodologies applicable in the development of new diagnostic tests
 CC for serotyping.
 SQ Sequence 14 AA;
 SQ 3 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
 SQ 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 1 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.27
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

X X
 GGQPLWITATK
 ||||
 VEFPLDITAGTEAA
 X 10

15. US-08-249-182-8 (1-11)

P98452 Sequence of C. trachomatis serovar H major outer m

ID P98452 standard; Protein; 14 AA.
 AC P98452;
 DT 06-MAR-1992 (first entry)
 DE Sequence of C. trachomatis serovar H major outer membrane protein (MOMP)
 DE variable domain (VD) H-VDIII encoded by base pairs 742-783
 KW Chlamydia trachomatis; antigen; monoclonal antibody; vaccine;
 KW diagnosis; serotyping; non-immunologic assay; ss.
 OS Chlamydia trachomatis.
 PN US7324664-A.
 PD 29-AUG-1989.
 PF 17-MAR-1989; 324664.
 PR 17-MAR-1989; US-324664.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Caldwell HD, Ying Y, Zhang YX, Watkins NG;
 DR WPI; 89-339697/46.
 DR N-PSDB; N97087.
 PT Chlamydia trachomatis genes - used for determn. of nucleotide and
 PT amino sequences of the variable domains of the major outer
 PT membrane proteins
 PS Disclosure; Fig 15; 49pp; English.
 CC The inventors sequenced the 4 MOMP VDs of ten C. trachomatis
 CC serovars and the amino acid sequences were deduced. The MOMP VDs
 CC with the greatest total hydrophilicity and charge values were found
 CC to be the location of antigenic determinants recognised by MOMP

specific monoclonal antibodies. The nucleotide, amino acid
 sequences and hydrophilicity/charge value analyses will assist in
 the selection of appropriate MOMP antigenic determinants to be used
 in the construction of synthetic peptides, subunits or recombinant
 chlamydial vaccines. This will allow the prodn. of reagents and
 methodologies applicable in the development of new diagnostic tests
 for serotyping.

Sequence 14 AA;
 4 A; 0 R; 0 N; 1 D; 0 B; 0 C; 0 Q; 2 E; 0 Z; 1 G; 0 H;
 1 I; 1 L; 0 K; 0 M; 1 F; 1 P; 0 S; 2 T; 0 W; 0 Y; 0 V;

Initial Score = 5 Optimized Score = 5 Significance = 3.27
 Residue Identity = 45% Matches = 5 Mismatches = 6
 Gaps = 0 Conservative Substitutions = 0

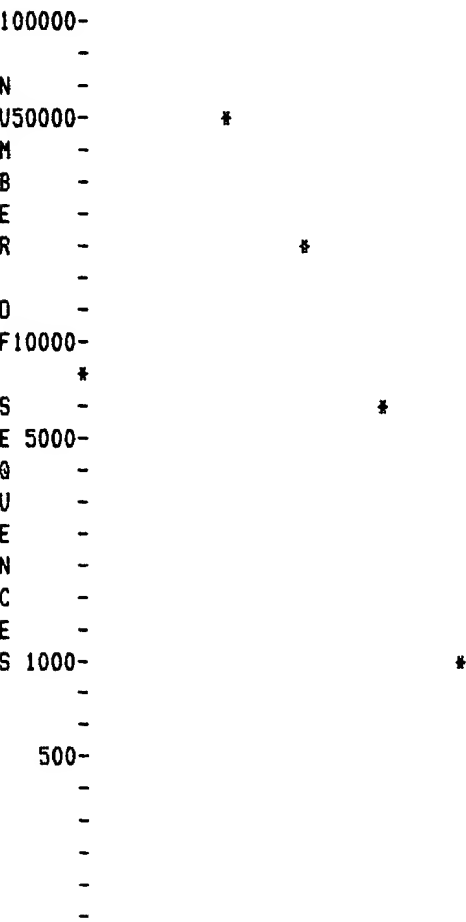
X X
 GGGPLWITATK
 ||||
 AEFPLDITAGTEAA
 X 10
 > 0 <
 0| 0 IntelliGenetics
 > 0 <

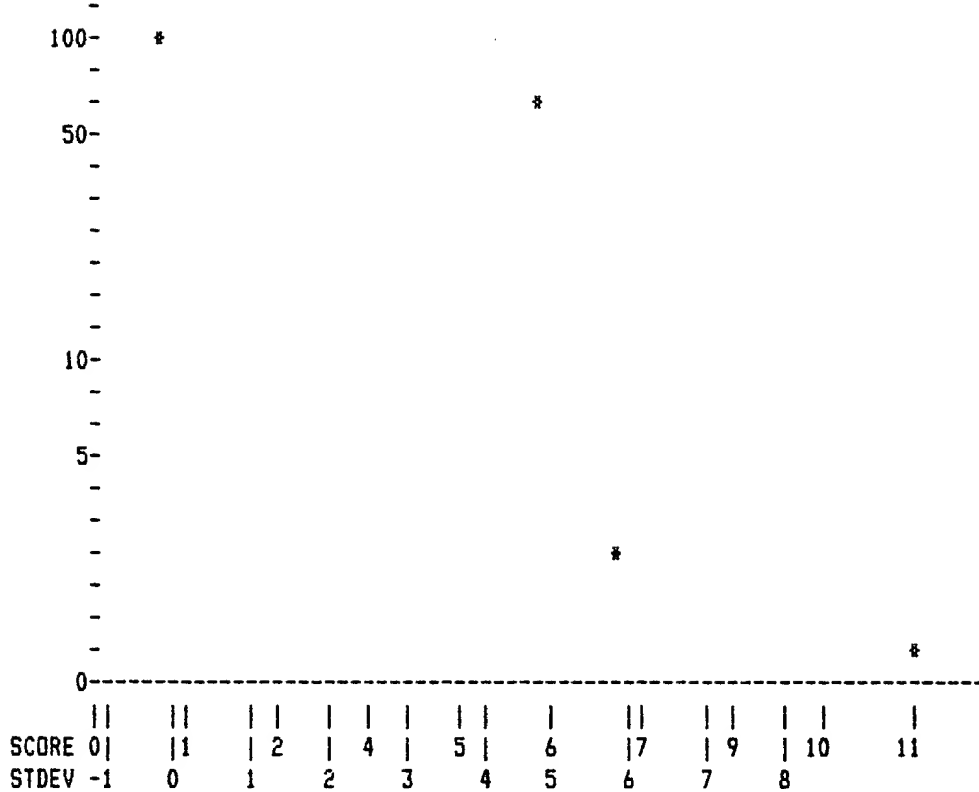
FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_8p.res made by on Thu 22 Sep 94 10:16:43-PDT.

Query sequence being compared:US-08-249-182-8 (1-11)
 Number of sequences searched: 70848
 Number of scores above cutoff: 4261

Results of the initial comparison of US-08-249-182-8 (1-11) with:
 Data bank : PIR 41, all entries





PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.98

Times:	CPU	Total Elapsed
	00:01:26.02	00:01:36.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	4261

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.
Cut-off raised to 5.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. A42329	autotaxin - human (fragments)	114	11	11	9.17	0

The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
2. WZBE2	gene 2 protein - human herpes	238	7	7	5.10	0
3. XYCHFA	fatty-acid synthase (EC 2.3.1	2446	7	7	5.10	0
**** 4 standard deviations above mean ****						
4. S20178	hypothetical protein (IFM1 3'	125	6	6	4.08	0
5. Y6ECC1	CFA1 fimbrial protein precurs	170	6	6	4.08	0
6. A38487	helix-destabilizing protein -	174	6	6	4.08	0
7. DDEC1B	helix-destabilizing protein -	175	6	6	4.08	0
8. B34078	prolactin-related protein III	213	6	6	4.08	0
9. S33428	spdA protein - Streptomyces a	224	6	6	4.08	0
10. C48652	transfer protein spdA - Strep	224	6	6	4.08	0
11. A44914	phosphate-dependent exoribonu	245	6	6	4.08	0
12. S03888	photosystem II oxygen-evolvin	248	6	6	4.08	0
13. S22763	photosystem II oxygen-evolvin	258	6	6	4.08	0
14. F2TOX2	photosystem II oxygen-evolvin	258	6	6	4.08	0
15. JS0771	photosystem II oxygen-evolvin	259	6	6	4.08	0
16. S07467	photosystem II oxygen-evolvin	259	6	6	4.08	0
17. S10016	photosystem II oxygen-evolvin	260	6	6	4.08	0
18. S15005	photosystem II oxygen-evolvin	265	6	6	4.08	0
19. IOHU3	insulin-like growth factor-bi	291	6	6	4.08	0
20. A36748	insulin-like growth factor-bi	292	6	6	4.08	0
21. MFNZRP	matrix protein - rinderpest v	335	6	6	4.08	0
22. S11738	hemagglutinin precursor - inf	378	6	7	4.08	0
23. A39314	gastricsin (EC 3.4.23.3) prec	384	6	6	4.08	0
24. S13094	glycerol-3-phosphate dehydrog	385	6	6	4.08	0
25. A31811	gastricsin (EC 3.4.23.3) prec	388	6	6	4.08	0
26. A29937	gastricsin (EC 3.4.23.3) prec	388	6	6	4.08	0
27. A38621	aspartate transaminase (EC 2.	392	6	6	4.08	0
28. A33510	gastricsin (EC 3.4.23.3) - ra	392	6	6	4.08	0
29. A24608	pepsin A (EC 3.4.23.1) precur	392	6	6	4.08	0
30. S29808	beta-3-adrenergic receptor -	400	6	6	4.08	0
31. A53281	beta 3-adrenergic receptor -	400	6	6	4.08	0
32. JS0349	hypothetical 45K protein (sbc	400	6	7	4.08	0
33. A41679	beta-3-adrenergic receptor -	400	6	6	4.08	0
34. A47041	transposase homolog (insertio	408	6	6	4.08	0
35. B39096	alkaline phosphatase (EC 3.1.	462	6	6	4.08	0
36. S16359	adenylate cyclase (EC 4.6.1.1	469	6	6	4.08	0
37. S41616	atpB protein - Euglena gracil	480	6	6	4.08	0
38. S34547	H+-transporting ATP synthase	480	6	6	4.08	0
39. A26951	nifB protein - Rhizobium meli	490	6	6	4.08	0
40. C24829	H+-transporting ATP synthase	491	6	6	4.08	0

1. US-08-249-182-8 (1-11)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
 Cioce, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis
 of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329

```

##status      preliminary
##molecule_type protein
##residues    1-114 ##label STR
##cross-references NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
                  NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
                  NCBIP:78509; NCBIP:78508; NCBIP:78503
##note        sequence extracted from NCBI backbone
SUMMARY       #length 114 #checksum 7335
SEQUENCE

```

```

Initial Score      =      11  Optimized Score =      11  Significance =  9.17
Residue Identity   =    100%  Matches          =      11  Mismatches   =      0
Gaps               =        0  Conservative Substitutions =      0

```

```

                                X      10
                                GG@PLWITATK
                                |||||
EFLSNYLTNVDDITLVPGLGRDIEHLTSLDFFRVNSM@TVFVVGYPGTFKGG@PLWITATKSPPFENINLYY
      10      20      30      40      50 X      60 X      70

DVPWNETIPEEVTXPNYL@AEVSYPAFKPXL DVYKWHVA
      80      90     100     110

```

2. US-08-249-182-8 (1-11)

WZBE2 gene 2 protein - human herpesvirus 3

```

ENTRY          WZBE2      #type complete
TITLE          gene 2 protein - human herpesvirus 3
ORGANISM       #formal_name human herpesvirus 3, varicella-zoster virus
DATE           30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change
                08-Apr-1994
ACCESSIONS     B27212
REFERENCE      A27345
#authors       Davison, A.J.; Scott, J.E.
#journal       J. Gen. Virol. (1986) 67:1759-1816
#title         The complete DNA sequence of varicella-zoster virus.
#cross-references MUID:86306657
#accession     B27212
##molecule_type DNA
##residues     1-238 ##label DAV

```

COMMENT The DNA sequence was obtained from EMBL, release 13.

GENETICS

```

#gene          2
CLASSIFICATION #superfamily varicella-zoster virus gene 2 protein
SUMMARY        #length 238 #molecular-weight 25984 #checksum 1948
SEQUENCE

```

```

Initial Score      =        7  Optimized Score =        7  Significance =  5.10
Residue Identity   =     63%  Matches          =         7  Mismatches   =         4
Gaps               =          0  Conservative Substitutions =          0

```

```

                                X      10
                                GG@PLWITATK
                                || | || |
SETLAYGHVPAFINGSTLVRPSLNATAEENPASETRCLLRVLAGRTVDLPGG@TLHITCTKTYVIICKYSKP
      10      20      30      40      50      X      60      X      70

GERLSLARLIGRAMTPGGARTFIILAMKEKRSTTLGYEC
      80      90     100     110

```

3. US-08-249-182-8 (1-11)

XYCHFA fatty-acid synthase (EC 2.3.1.85) - chicken

```

ENTRY          XYCHFA      #type complete

```

TITLE fatty-acid synthase (EC 2.3.1.85) - chicken
 CONTAINS 3-hydroxypalmitoyl-[acyl-carrier-protein] dehydratase (EC 4.2.1.61); 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100); 3-oxoacyl-[acyl-carrier-protein] synthase (EC 2.3.1.41); enoyl-[acyl-carrier-protein] reductase (NADPH, B-specific) (EC 1.3.1.10); oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14); [acyl-carrier-protein] S-acetyltransferase (EC 2.3.1.38); [acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39)

ORGANISM #formal_name Gallus gallus #common_name chicken
 DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 31-Dec-1993

ACCESSIONS A33918; A30445; A31236; B31236; A30297; A31184; A31185; A30446; S03856

REFERENCE A33918
 #authors Holzer, K.P.; Liu, W.; Hammes, G.G.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:4387-4391
 #title Molecular cloning and sequencing of chicken liver fatty acid synthase cDNA.
 #cross-references MUID:89282777
 #accession A33918
 ##molecule_type mRNA
 ##residues 1-1701 ##label HOL
 #accession A30445
 ##molecule_type protein
 ##residues 33-39;1012-1017 ##label HOL2

REFERENCE A31236
 #authors Yuan, Z.; Liu, W.; Hammes, G.G.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1988) 85:6328-6331
 #title Molecular cloning and sequencing of DNA complementary to chicken liver fatty acid synthase mRNA.
 #cross-references MUID:88320436
 #accession A31236
 ##molecule_type mRNA
 ##residues 1678-2446 ##label YUA
 #accession B31236
 ##molecule_type mRNA
 ##residues 1678-2275,2284-2446 ##label YUA2

REFERENCE A30297
 #authors Chirala, S.S.; Kasturi, R.; Pazirandeh, M.; Stolow, D.T.; Huang, W.Y.; Wakil, S.J.
 #journal J. Biol. Chem. (1989) 264:3750-3757
 #title A novel cDNA extension procedure. Isolation of chicken fatty acid synthase cDNA clones.
 #cross-references MUID:89139426
 #accession A30297
 ##molecule_type mRNA
 ##residues 1494-1503,'W',1505-1658,'Q',1660-1671,'S',1673-2275,2284-2446 ##label CH1
 ##cross-references ENBL:J04485

REFERENCE A31184
 #authors Yang, C.Y.; Huang, W.Y.; Chirala, S.; Wakil, S.J.
 #journal Biochemistry (1988) 27:7773-7777
 #title Complete amino acid sequence of the thioesterase domain of chicken liver fatty acid synthase.
 #cross-references MUID:89088151
 #accession A31184
 ##molecule_type protein
 ##residues 2135-2275,2284-2442 ##label YAN

REFERENCE A31185
 #authors Kasturi, R.; Chirala, S.; Pazirandeh, M.; Wakil, S.J.
 #journal Biochemistry (1988) 27:7778-7785
 #title Characterization of a genomic and cDNA clone coding for the thioesterase domain and 3' noncoding region of the chicken liver fatty acid synthase gene.
 #cross-references MUID:89088152


```

#accession      A31183
##molecule_type DNA
##residues      2128-2275,2284-2446 ##label KAS
##cross-references EMBL:J02839
#accession      A30446
##molecule_type mRNA
##residues      2128-2275,2284-2446 ##label KA2
REFERENCE      S03856
#authors        Huang, W.Y.; Stoops, J.K.; Wakil, S.J.
#journal         Arch. Biochem. Biophys. (1989) 270:92-98
#title           Complete amino acid sequence of chicken liver acyl carrier
                  protein derived from the fatty acid synthase.
#cross-references MUID:89192401
#accession      S03856
##molecule_type protein
##residues      2047-2135 ##label HUA
CLASSIFICATION #superfamily rat fatty-acid synthase
KEYWORDS         acyltransferase; carbon-oxygen lyase; carrier protein; fatty
                  acid biosynthesis; homodimer; hydro-lyase; mammary gland;
                  oxidoreductase; phosphopantetheine; thiolester hydrolase
FEATURE
  1-2446          #protein fatty-acid synthase (isoform 1) #status
                  predicted #label MAT\
  1-2275,2284-2446 #protein fatty-acid synthase (isoform 2) #status
                  predicted #label MA2\
  1-338           #domain 3-oxoacyl-[acyl-carrier-protein] synthase
                  #status predicted #label ENZ1\
  1248-1266       #region active-site region (of
                  3-hydroxypalmitoyl-[acyl-carrier-protein] dehydratase)
                  #status predicted\
  2047-2143       #domain acyl carrier protein #status predicted #label
                  ACP\
  2144-2446       #domain oleoyl-[acyl-carrier-protein] hydrolase #status
                  predicted #label ENZ7\
  87              #active_site Cys (of 3-oxoacyl-[acyl-carrier-protein]
                  synthase) #status predicted\
  506             #active_site Ser (of [acyl-carrier-protein]
                  acetyl/malonyltransferase) #status predicted\
  1631            #active_site Ser (of enoyl-[acyl-carrier-protein]
                  reductase) #status predicted\
  1634            #active_site Lys (of enoyl-[acyl-carrier-protein]
                  reductase) #status predicted\
  1856            #active_site Lys (of 3-oxoacyl-[acyl-carrier-protein]
                  reductase) #status predicted\
  2084            #binding_site phosphopantetheine (Ser) (covalent) (in
                  acyl carrier protein) #status predicted\
  2235            #active_site Ser (of oleoyl-[acyl-carrier-protein]
                  hydrolase) #status predicted
SUMMARY          #length 2446 #molecular-weight 267247 #checksum 8579
SEQUENCE

```

```

Initial Score    =      7  Optimized Score =      7  Significance =  5.10
Residue Identity =   63%  Matches           =      7  Mismatches   =    4
Gaps             =      0  Conservative Substitutions =      0

```

```

      X      X
      GGQPLWITATK
      ||| |||

```

```

LNLVAMKRSFFGVSIFLCRRQSPAKAPILLPVDDTHYKWVDSLKEILADSSEQPLWLTATNCGNSGILGMVN
1320      1330      1340      1350      1360      1370      1380

```

```

CLRLEAEGHRIRC VFVSNLSPSSTVPATSLSSLEMQKII
1390      1400      1410      1420

```

S20178 hypothetical protein (IFM1 3' region) - yeast (Sac

ENTRY S20178 #type fragment
TITLE hypothetical protein (IFM1 3' region) - yeast (Saccharomyces cerevisiae) (fragment)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 30-Sep-1993
ACCESSIONS S20178; S17024
REFERENCE S20177
#authors Vambutas, A.; Ackerman, S.H.; Tzagoloff, A.
#journal Eur. J. Biochem. (1991) 201:643-652
#title Mitochondrial translational-initiation and elongation factors in Saccharomyces cerevisiae.
#cross-references MUID:92037620
#accession S20178
##molecule_type DNA
##residues 1-125 ##label VAM
##cross-references EMBL:X58379
SUMMARY #length 125 #checksum 772
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10
GGGPLWITATK
|| || ||
PRTEKLSKFLDDDTFQKFQEVVGYNPLQVLRDYGKPLLYAETKVDILSTVPRPGYNPSSQRIFEMQLMPK
10 20 30 X 40 X 50 60 70

MIFDLEEVVSVDNGMEWGILVF
80 90

5. US-08-249-182-8 (1-11)

YQECC1 CFA1 fimbrial protein precursor - Escherichia coli

ENTRY YQECC1 #type complete
TITLE CFA1 fimbrial protein precursor - Escherichia coli
ALTERNATE_NAMES CFA1 pilin; colonization factor antigen I (CFA1)
ORGANISM #formal_name Escherichia coli
DATE 14-Nov-1983 #sequence_revision 30-Jun-1991 #text_change 31-Dec-1993
ACCESSIONS A30589; A03495; A43831
REFERENCE A30589
#authors Karjalainen, T.K.; Evans, D.G.; So, M.; Lee, C.H.
#journal Infect. Immun. (1989) 57:1126-1130
#title Molecular cloning and nucleotide sequence of the colonization factor antigen I gene of Escherichia coli.
#cross-references MUID:89173309
#accession A30589
##molecule_type DNA
##residues 1-170 ##label KAR
REFERENCE A03495
#authors Klemm, P.
#journal Eur. J. Biochem. (1982) 124:339-348
#title Primary structure of the CFA1 fimbrial protein from human enterotoxigenic Escherichia coli strains.
#cross-references MUID:82235736
#contents Strain H10407
#accession A03495
##molecule_type DNA
##residues 24-75,'N',77-96,'A',98-170 ##label KLE
REFERENCE A43831

#authors Cassels, P.J.; Deal, C.D.; Reid, R.H.; Jarboe, D.L.; Nauss, J.L.; Carter, J.M.; Boedeker, E.C.
#journal Infect. Immun. (1992) 60:2174-2181
#title Analysis of Escherichia coli colonization factor antigen I linear B-cell epitopes, as determined by primate responses, following protein sequence verification.
#cross-references MUID:92267624
#contents Strain H10407
#accession A43831
##molecule_type protein
##residues 24-170 ##label CAS
##cross-references NCBIP:104220
##note sequence extracted from NCBI backbone
COMMENT The CFA1 fimbriae are rather rigid, thread-like filaments of 0.5-1 micrometer, with an apparent axial hole, and a diameter of 7 nanometers. A single CFA1 fimbria consists of about 100 identical protein subunits.
CLASSIFICATION #superfamily CFA1 fimbrial protein
KEYWORDS fimbria; fimbrial antigen
FEATURE
1-23 #domain signal sequence #label SIG\
24-170 #protein CFA1 fimbrial protein #label MAT
SUMMARY #length 170 #molecular-weight 17461 #checksum 1166
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X X
GGQPLWITATK
||| | ||
PASKTFESYRVMTQVHTNDATKKVIVKLDTPQLTDVLNSTVQMPISVSWGQVLSTAKEFEAAALGYSAS
60 70 80 90 100 110 120
GVNGVSSSQELVISAAPKTAGTAPTAGNYSGVVSLVMTL
130 140 150 160

6. US-08-249-182-8 (1-11)

A38487 helix-destabilizing protein - Escherichia coli pla

ENTRY A38487 #type fragment
TITLE helix-destabilizing protein - Escherichia coli plasmid R64 (fragment)
ORGANISM #formal_name Escherichia coli
DATE 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 18-Jun-1993
ACCESSIONS A38487
REFERENCE A38487
#authors Ruvolo, P.P.; Keating, K.M.; Williams, K.R.; Chase, J.W.
#journal Proteins (1991) 9:120-134
#title Single-stranded DNA binding proteins (SSBs) from prokaryotic transmissible plasmids.
#cross-references MUID:91180109
#accession A38487
##status preliminary
##molecule_type DNA
##residues 1-174 ##label RUV
##note sequence not compared to nucleotide translation
GENETICS
#genome plasmid
CLASSIFICATION #superfamily bacterial helix-destabilizing protein
SUMMARY #length 174 #checksum 691
SEQUENCE

ORGANISM #formal_name Bos primigenius laurus #common_name cattle
 DATE 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change
 30-Sep-1993
 ACCESSIONS B34078
 REFERENCE A34078
 #authors Kessler, M.A.; Milosavljevic, M.; Zieler, C.G.; Schuler, L.A.
 #journal Biochemistry (1989) 28:5154-5161
 #title A subfamily of bovine prolactin-related transcripts distinct
 from placental lactogen in the fetal placenta.
 #cross-references MUID:89352599
 #accession B34078
 ##status preliminary
 ##molecule_type mRNA
 ##residues 1-213 ##label KES
 ##cross-references GB:M27240
 SUMMARY #length 213 #molecular-weight 25016 #checksum 9382
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                || | |||
QVNSCPSCCPDVFDIPLESLTHLFLNASRLSHDIVNHTTINFHEFDEKYAQN@PYTINATKSCHTNSLHTPG
10      20      30      40      50      60      70      80

EREKALRMNEDLSKWILMLLYSWHRPLYLLVKDLQSMK
      90      100      110      120
  
```

9. US-08-249-182-8 (1-11)

S33428 spdA protein - Streptomyces ambofaciens

ENTRY S33428 #type complete
 TITLE spdA protein - Streptomyces ambofaciens
 ORGANISM #formal_name Streptomyces ambofaciens
 DATE 06-Mar-1994; #sequence_revision 06-Mar-1994; #text_change
 06-Mar-1994
 ACCESSIONS S33428
 REFERENCE S33420
 #authors Hagege, J.M.; Pernodet, J.L.; Gerbaud, C.; Sezonov, G.;
 Friedmann, A.; Guerineau, M.
 #submission submitted to the EMBL Data Library, January 1993
 #accession S33428
 ##status preliminary
 ##residues 1-224 ##label HAG
 ##cross-references EMBL:Z19593
 SUMMARY #length 224 #molecular-weight 23575 #checksum 9089
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                GG@PLWITATK
                                || | |||
GPSRLAWSWFVIALVASLGANVATAGLLDLNDVPAWLRILVAAMPALAFMGGTLLAHTATHHEPEAPAPT@P
      70      80      90      100      110 X      120 X      130

APEPPAFTDEHDLVRVDDTEEPPELPAPGLQ@APAPPAV
      140      150      160      170
  
```

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                |||    |||
YIEG@LRTRSWDDNGITRYITEILVKTTGTMTQMLGSAP@QNAQA@PKP@QNG@P@SADATKKGGAKTKGRER
 80      90      100      110      120      130      140

KAAQPEP@P@TPEGEDYGFSDDIPF
150      160      170

```

7. US-08-249-182-8 (1-11)

DDECIB helix-destabilizing protein - Escherichia coli pla

ENTRY DDECIB #type complete
 TITLE helix-destabilizing protein - Escherichia coli plasmid
 Collb-P9
 ALTERNATE_NAMES single-stranded DNA-binding protein
 ORGANISM #formal_name Escherichia coli
 DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
 30-Jun-1993
 ACCESSIONS A32304
 REFERENCE A32304
 #authors Howland, C.J.; Rees, C.E.D.; Barth, P.T.; Wilkins, B.M.
 #journal J. Bacteriol. (1989) 171:2466-2473
 #title The ssb gene of plasmid Collb-P9.
 #cross-references MUID:89213928
 #accession A32304
 ##molecule_type DNA
 ##residues 1-175 ##label HOW
 COMMENT The plasmid-encoded single-stranded DNA-binding proteins may be
 involved in DNA metabolism during bacterial conjugation; their
 functions are also related to the plasmid-mediated SOS initiation
 process.
 GENETICS
 #gene ssb
 #genome plasmid
 CLASSIFICATION #superfamily bacterial helix-destabilizing protein
 KEYWORDS DNA repair; DNA replication; single-stranded DNA binding
 FEATURE
 2-174 #protein helix-destabilizing protein #label MAT
 SUMMARY #length 175 #molecular-weight 19240 #checksum 1668
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                |||    |||
YIEG@LRTRSWDDNGITRYITEILVKTTGTMTQMLGSAP@QNAQA@PKP@QNG@P@SADATKKGGAKTKGRGR
 80      90      100      110      120      130      140      150

KAAQPEP@P@TPEGEDYGFSDDIPF
      160      170

```

8. US-08-249-182-8 (1-11)

B34078 prolactin-related protein III - bovine

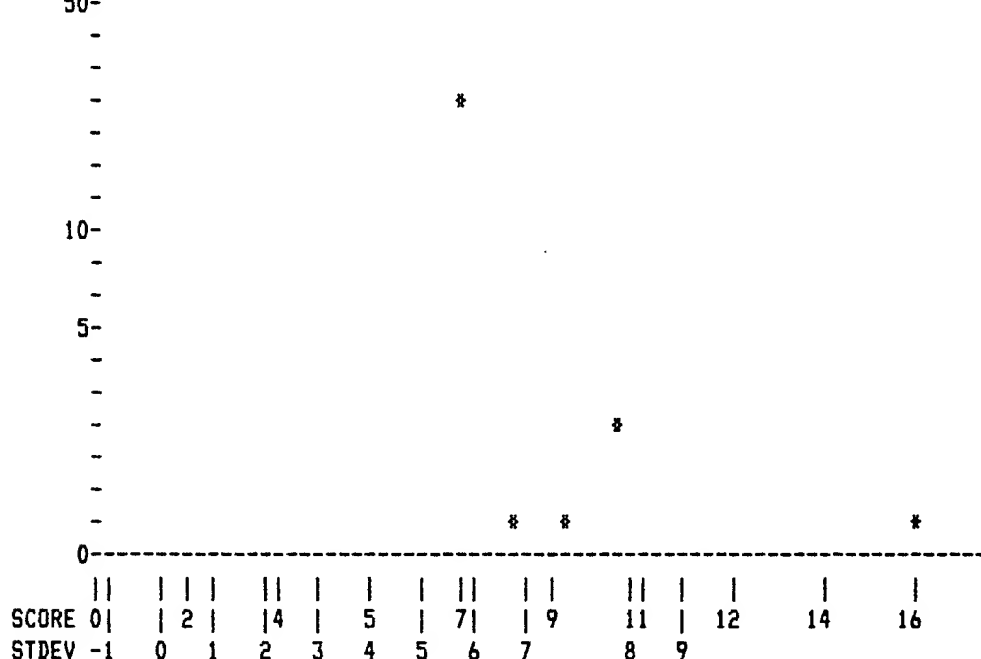
ENTRY B34078 #type complete
 TITLE prolactin-related protein III - bovine

Sequence Name	Description	Length	Score	Sig.	Frame
**** 7 standard deviations above mean ****					
2. A39216	plasma cell membrane protein	925	10	11	7.48 0
3. S21706	nucleotide pyrophosphatase -	925	10	11	7.48 0
**** 6 standard deviations above mean ****					
4. A27410	plasma cell membrane protein	905	9	9	6.55 0
**** 5 standard deviations above mean ****					
5. RKAALC	ribulose-bisphosphate carboxy	474	8	8	5.61 0
**** 4 standard deviations above mean ****					
6. DSNVAC	superoxide dismutase (EC 1.15	151	7	7	4.68 0
7. A35216	FPD4 protein - fowlpox virus	218	7	8	4.68 0
8. A24637	T-cell surface glycoprotein C	236	7	7	4.68 0
9. A45442	Sec13p=endoplasmic reticulum	297	7	7	4.68 0
10. S30803	SEC13 protein - yeast (Saccha	297	7	7	4.68 0
11. S06938	glutamate dehydrogenase (NADP	446	7	7	4.68 0
12. A33504	glutamate dehydrogenase (EC 1	447	7	7	4.68 0
13. A22413	glutamate dehydrogenase (NADP	447	7	7	4.68 0
14. DEECEN	glutamate dehydrogenase (NADP	447	7	7	4.68 0
15. APBOL	leucyl aminopeptidase (EC 3.4	478	7	7	4.68 0
16. A61191	nuclear hormone receptor ST-5	564	7	7	4.68 0
17. HNNZB3	hemagglutinin-neuraminidase (572	7	7	4.68 0
18. S28142	SEC23 protein homolog, 64.7K	573	7	7	4.68 0
19. A37251	early response protein NAK1 -	598	7	7	4.68 0
20. J01894	gene p74 protein - pigeonpox	630	7	7	4.68 0
21. A30347	exotoxin A precursor - Pseudo	638	7	7	4.68 0
22. QRECFC	ferrienterochelin receptor pr	745	7	7	4.68 0
23. H36790	hypothetical protein ORF43 -	891	7	7	4.68 0
24. S07421	E2 glycoprotein precursor - a	1162	7	7	4.68 0
25. VGIHAK	E2 glycoprotein precursor - a	1162	7	7	4.68 0
26. S12127	protein-tyrosine kinase (EC 2	1187	7	7	4.68 0
27. J01258	RNA-directed RNA polymerase (1335	7	7	4.68 0
28. S32437	polyprotein - Volvox carteri	1462	7	9	4.68 0
29. A30788	mannose 6-phosphate receptor	2499	7	7	4.68 0
**** 3 standard deviations above mean ****					
30. S27205	interleukin-2 precursor - mou	63	6	6	3.74 0
31. S00972	kcrB2 protein - Escherichia c	66	6	6	3.74 0
32. PN0088	matrix protein M3 - influenza	68	6	7	3.74 0
33. PN0085	matrix protein M3 - influenza	68	6	7	3.74 0
34. E42516	D-ORF-B protein - vaccinia vi	80	6	6	3.74 0
35. Q0V24	hypothetical protein B-80 - v	80	6	6	3.74 0
36. S28564	hypothetical protein 2 - phag	92	6	6	3.74 0
37. S40170	gene fdxB protein - Plectonem	96	6	7	3.74 0
38. S10703	plastocyanin - parsley	97	6	7	3.74 0
39. S10702	plastocyanin - parsley	97	6	7	3.74 0
40. S06105	plastocyanin - rice	97	6	6	3.74 0

1. US-08-249-182-9 (1-16)

A42329 autotaxin - human (fragments)

ENTRY A42329 #type fragments
 TITLE autotaxin - human (fragments)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 04-Mar-1993; #sequence_revision 01-Jan-1993; #text_change
 08-May-1993
 ACCESSIONS A42329
 REFERENCE A42329
 #authors Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.;
 Ciocer, V.; Schiffmann, E.; Liotta, L.A.
 #journal J. Biol. Chem. (1992) 267:2524-2529
 #title Identification, purification, and partial sequence analysis
 of autotaxin, a novel motility-stimulating protein.
 #cross-references MUID:92129337
 #accession A42329



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	1.07

Times:	CPU	Total Elapsed
	00:01:27.02	00:01:35.00

Number of residues:	20816057
Number of sequences searched:	70848
Number of scores above cutoff:	3804

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.
 Cut-off raised to 6.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. A42329	autotaxin - human (fragments)	114	16	16	13.09	0

The list of other best scores is:

Sequence 420 AA;
 SQ 57 A; 32 R; 9 N; 24 D; 0 B; 8 C; 21 Q; 32 E; 0 Z; 41 G; 11 H;
 SQ 13 I; 43 L; 4 K; 2 M; 12 F; 28 P; 22 S; 19 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSMOTVFVGYGPTFK
 ||||| ||
 CAGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNWTVRLLGAHRQLEERGVYFVGYHGTFFLEAAQSI
 190 200 210 220 230 X 240 250

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYA@DQEPDARGRIR
 260 270 280 290 300

> O <
 O| |O IntelliGenetics
 > O <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file u249_9p.res made by on Thu 22 Sep 94 10:23:45-PDT.

Query sequence being compared:US-08-249-182-9 (1-16)
 Number of sequences searched: 70848
 Number of scores above cutoff: 3804

Results of the initial comparison of US-08-249-182-9 (1-16) with:
 Data bank : PIR 41, all entries

100000-
 -
 N -
 U50000-
 M - *
 B -
 E -
 R - *
 -
 O -
 F10000- *
 -
 S *
 E 5000-
 Q -
 U -
 E -
 N - *
 C -
 E -
 S 1000-
 -
 -
 500-
 -
 -
 -
 -
 -
 100- *
 -
 -

AC R06974;
 DT 18-JAN-1991 (first entry)
 DE PE40ab protein comprising a portion of the Pseudomonas exotoxin A
 DE lacking cysteine residues at 265, 287, 372 and 379.
 KW TGF-alpha-PE40; PE40ab; tumour; epidermal growth factor; EGF;
 KW transforming growth factor-alpha; TGF-alpha.
 OS Pseudomonas sp.
 PN EP-389043-A.
 PD 26-SEP-1990.
 PF 15-MAR-1990; 200613.
 PR 22-MAR-1989; US-327214.
 PA (MERI) MERCK & CO INC.
 PI Riemen MW, Stirdivant SM;
 DR WPI; 90-291988/39.
 PT Modified PE40 by substitution with other amino acids for cysteine -
 PT improving specificity of targetting agent for tumour cells.
 PS Disclosure; Table 6; 21pp; English.
 CC By replacing cysteine residues at positions 265 and 287 and/or 372
 CC and 379, chemical ambiguities may be eliminated, and targeting
 CC specificity for targeted agents of tumour cells eg. EGF or TGF-alpha
 CC may be improved.
 CC See also 006127.
 SQ Sequence 420 AA;
 SQ 59 A; 32 R; 9 N; 24 D; 0 B; 6 C; 21 Q; 31 E; 0 Z; 42 G; 11 H;
 SQ 13 I; 43 L; 4 K; 2 M; 12 F; 28 P; 21 S; 20 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                |||||  ||
AAGPADSGDALLERNYPTGA EFLGDGGDVSFSTRGTQNWTV ERL LQAHRQLEERG YV F V G Y H G T F L E A A Q S I
 190      200      210      220      230      X 240      250

VFGGVRRARSQDLDAIW RGFYIAGDPALAYGYA QDQEPDARGRIR
260      270      280      290      300

```

15. US-08-249-182-9 (1-16)
 R06992 PE40aB protein comprising a portion of the Pseudon

ID R06992 standard; protein; 420 AA.
 AC R06992;
 DT 18-JAN-1991 (first entry)
 DE PE40aB protein comprising a portion of the Pseudomonas exotoxin A
 DE lacking cysteine residues at 265 and 287.
 KW TGF-alpha-PE40; PE40ab; tumour; epidermal growth factor; EGF;
 KW transforming growth factor-alpha; TGF-alpha.
 OS Pseudomonas sp.
 PN EP-389043-A.
 PD 26-SEP-1990.
 PF 15-MAR-1990; 200613.
 PR 22-MAR-1989; US-327214.
 PA (MERI) MERCK & CO INC.
 PI Riemen MW, Stirdivant SM;
 DR WPI; 90-291988/39.
 PT Modified PE40 by substitution with other amino acids for cysteine -
 PT improving specificity of targetting agent for tumour cells.
 PS Disclosure; Table 4; 21pp; English.
 CC By replacing cysteine residues at positions 265 and 287 and/or 372
 CC and 379, chemical ambiguities may be eliminated, and targeting
 CC specificity for targeted agents of tumour cells eg. EGF or TGF-alpha
 CC may be improved.
 CC See also 006127.

CAGPADSGDALLERNYPTGAFLGDDGVSFSTRGTQNWTVRLLQAHRLQLEERGTVFVGYHGTFLAAQSI
 190 200 210 220 230 X 240 250

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
 260 270 280 290 300

13. US-08-249-182-9 (1-16)

R06447 TGF-57-Pseudomonas exotoxin 40 fusion protein.

ID R06447 standard; protein; 420 AA.
 AC R06447;
 DT 04-JAN-1991 (first entry)
 DE TGF-57-Pseudomonas exotoxin 40 fusion protein.
 KW Pseudomonas exotoxin-40 (PE40); protein targeting agent;
 KW psoriasis treatment; anti-tumour agent;
 FH Key Location/Qualifiers
 FT Region 1..54
 FT /label=residues -4 to +50 of TGF-alpha
 FT Region 58..420
 FT /label=residues +252 to +613 of PE
 PN EP-383599-A.
 PD 22-AUG-1990.
 PF 15-FEB-1990; 301639.
 PR 17-FEB-1989; US-312540.
 PR 03-AUG-1989; US-389092.
 PR 21-DEC-1989; US-449187.
 PA (MERI) MERCK & CO INC.
 PI Oliff A, Jones DD, Edwards GM;.
 DR WPI; 90-255832/34.
 DR N-PSDB; 005666.
 PT Modified pseudomonas exotoxin hybrid proteins - has at least 2
 PT cysteine residues replaced or deleted to improve binding to
 PT receptors.
 PS Example ; Table 2; 20pp; English.
 CC Modified pseudomonas exotoxin (PE40) linked to
 CC 5' portion of transforming growth factor (TGF)-alpha as a targeting
 CC agent. The corresponding nucleotide sequence was constructed from
 CC a synthetic oligonucleotide encoding the 5' portion of TGF-alpha
 CC and linked to PE40 and a linker cassette called "cassette 57". The
 CC recombinant plasmid was used to transform E.coli JM109 cells.
 CC The hybrid protein can bind and kill tumour cells or keratinocytes
 CC possessing TGF receptors for treatment of psoriasis or warts.
 CC See also R06448-R06450
 SQ Sequence 420 AA;
 SQ 55 A; 32 R; 9 N; 24 D; 0 B; 10 C; 21 Q; 31 E; 0 Z; 42 G; 11 H;
 SQ 13 I; 42 L; 4 K; 2 M; 13 F; 28 P; 22 S; 19 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNMQTVFVGYGPTFK
 ||||| ||
 CAGPADSGDALLERNYPTGAFLGDDGVSFSTRGTQNWTVRLLQAHRLQLEERGTVFVGYHGTFLAAQSI
 190 200 210 220 230 X 240 250
 VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
 260 270 280 290 300

14. US-08-249-182-9 (1-16)

R06994 PE40ab protein comprising a portion of the Pseudom

ID R06994 standard; protein; 420 AA.

Sequence 420 AA;
SQ 58 A; 32 R; 9 N; 24 D; 0 B; 8 C; 21 Q; 30 E; 0 Z; 43 G; 11 H;
SQ 12 I; 42 L; 4 K; 2 M; 13 F; 28 P; 21 S; 20 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||||| ||
AAGPADSGDALLERNYPTGAFLDGGDVSFSTRGTQNWTVRLLQAHRLGLEERGYYFVGYHGTFLEAAQSI
  190      200      210      220      230      X 240      250

VFGGVRARSQDLDAIWRFYIAGDPALAYGYAQQDQEPDARGRIR
  260      270      280      290      300
```

12. US-08-249-182-9 (1-16)

R06448 TGF-alpha-PE40-aB modified pseudomonas exotoxin hy

ID R06448 standard; protein; 420 AA.
AC R06448;
DT 04-JAN-1991 (first entry)
DE TGF-alpha-PE40-aB modified pseudomonas exotoxin hybrid protein.
KW Pseudomonas exotoxin-40 (PE40); protein targeting agent;
KW psoriasis treatment; anti-tumour agent; TGF-alpha-PE40-aB;
FH Key Location/Qualifiers
FT Region 1..54
FT /label=residues -4 to +50 of TGF-alpha
FT Region 58..420
FT /label=residues +252 to +613 of PE
FT /note="residues 265, 272 and 287 are modified"
PN EP-383599-A.
PD 22-AUG-1990.
PF 15-FEB-1990; 301639.
PR 17-FEB-1989; US-312540.
PR 03-AUG-1989; US-389092.
PR 21-DEC-1989; US-449187.
PA (MERI) MERCK & CO INC.
PI Oliff A, Jones DD, Edwards GM;.
DR WPI; 90-255832/34.
PT Modified pseudomonas exotoxin hybrid proteins - has at least 2
PT cysteine residues replaced or deleted to improve binding to
PT receptors.
PS Example ; Table 4; 20pp; English.
CC Modified pseudomonas exotoxin (PE40) linked to
CC 5' portion of transforming growth factor (TGF)-alpha as a targeting
CC agent. Three site-specific mutations have been introduced c.f wild-
CC type PE40. The Cys residues at positions 265 and 287 of the
CC exotoxin have been replaced by Ala residues. Phe at position 272
CC has been replaced with Leu. These changes improve receptor binding.
CC The hybrid protein can bind and kill tumour cells or keratinocytes
CC possessing TGF receptors for treatment of psoriasis or warts.
CC See also R06447 and R06449-R06450
SQ Sequence 420 AA;
SQ 57 A; 32 R; 9 N; 24 D; 0 B; 8 C; 21 Q; 32 E; 0 Z; 41 G; 11 H;
SQ 13 I; 43 L; 4 K; 2 M; 12 F; 28 P; 22 S; 19 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||||| ||
```


PI Ahern J, Helmsbrook DC, Oliff A, Stirdivant SM;
 DR WPI; 92-026359/04.
 PT Treatment of bladder cancer using hybrid protein - comprising
 PT cell targetting agent e.g. EGF that binds to EGF receptor on
 PT tumour cells and PE40 cell toxin
 PS Disclosure; Page 18-19; 34pp; English.
 CC The modified PE40 domains of the hybrid proteins have two or four of
 CC the Cys residues (designated Cys265, Cys287, Cys372 and Cys372)
 CC substituted with neutral amino acids, e.g. Gly, Ala, or Phe.
 CC TGF-alpha-PE40aB (R20199) has Cys265 and Cys287 replaced;
 CC TGF-alpha-PE40Ab (R20200) has Cys372 and Cys379 replaced; and
 CC TGF-alpha-PE40ab (R20201) has all four replaced.
 CC The modified hybrid proteins were produced in E.coli transformed
 CC with TAC expression vectors. Site specific mutations were introduced
 CC to the unmodified TGF-alpha-PE40 gene cloned in pTACTGF57-PE40.
 CC The mol. efficiently targets receptors on human bladder tumour cells
 CC (the modified PE40 domain has improved receptor binding) and is
 CC used for selectively killing bladder tumour cells.
 SQ Sequence 420 AA;
 SQ 57 A; 32 R; 9 N; 24 D; 0 B; 8 C; 22 Q; 30 E; 0 Z; 42 G; 11 H;
 SQ 13 I; 42 L; 4 K; 2 M; 13 F; 28 P; 21 S; 20 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQTVFVGYGPTFK
                                ||||| ||
AAGPADSGDALLERNYPTGAFLGDGGDVSFSTRGTQNWTVRLLQAHRLQLEERGYVFGYHGTFLEAAQSI
  190      200      210      220      230      X 240      250

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
  260      270      280      290      300

```

10. US-08-249-182-9 (1-16)

R06450 TGF-alpha-PE40-ab modified pseudomonas exotoxin hy

ID R06450 standard; protein; 420 AA.
 AC R06450;
 DT 04-JAN-1991 (first entry)
 DE TGF-alpha-PE40-ab modified pseudomonas exotoxin hybrid protein.
 KW Pseudomonas exotoxin-40 (PE40); protein targeting agent;
 KW psoriasis treatment; anti-tumour agent; TGF-alpha-PE40-ab;
 FH Key Location/Qualifiers
 FT Region 1..54
 FT /label=residues -4 to +50 of TGF-alpha
 FT Region 58..420
 FT /label=residues +252 to +613 of PE
 FT /note="residues 265, 272, 287, 369, 372 and 379
 FT are modified"
 PN EP-383599-A.
 PD 22-AUG-1990.
 PF 15-FEB-1990; 301639.
 PR 17-FEB-1989; US-312540.
 PR 03-AUG-1989; US-389092.
 PR 21-DEC-1989; US-449187.
 PA (MERI) MERCK & CO INC.
 PI Oliff A, Jones DD, Edwards GM;
 DR WPI; 90-255832/34.
 PT Modified pseudomonas exotoxin hybrid proteins - has at least 2
 PT cysteine residues replaced or deleted to improve binding to
 PT receptors.
 PS Example ; Table 6; 20pp; English.
 CC Modified pseudomonas exotoxin (PE40) linked to

ID R07054 standard; protein; 419 AA.
AC R07054;
DT 18-JAN-1991 (first entry)
DE PE40AB protein comprising a portion of the Pseudomonas exotoxin A.
KW TGF-alpha-PE40; PE40ab; tumour; epidermal growth factor; EGF;
KW transforming growth factor-alpha; TGF-alpha.
OS Pseudomonas sp.
PN EP-389043-A.
PD 26-SEP-1990.
PF 15-MAR-1990; 200613.
PR 22-MAR-1989; US-327214.
PA (MERI) MERCK & CO INC.
PI Riemen MW, Stirdivant SM;
DR WPI; 90-291988/39.
DR N-PSDB; 006127.
PT Modified PE40 by substitution with other amino acids for cysteine -
PT improving specificity of targetting agent for tumour cells.
PS Disclosure; Table 3; 21pp; English.
CC By replacing cysteine residues at positions 265 and 287 and/or 372
CC and 379, chemical ambiguities may be eliminated, and targeting
CC specificity for targeted agents of tumour cells eg. EGF or TGF-alpha
CC may be improved.
SQ Sequence 419 AA;
SQ 55 A; 31 R; 9 N; 24 D; 0 B; 10 C; 20 Q; 32 E; 0 Z; 42 G; 11 H;
SQ 13 I; 42 L; 4 K; 2 M; 13 F; 28 P; 22 S; 19 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

X
10
X

VNSMGTVFVGYGPTFK

||||| ||

CAGPADSGDALLERNYPTGAFLGDDGDSFSFSTRGTQNWTVRLLQAHRELEERGYYFVGYHGTFLCAAQSI

190200210220230X240250

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR

260270280290300

9. US-08-249-182-9 (1-16)
R20200 TGF-alpha-PE40Ab.

ID R20200 standard; Protein; 420 AA.
AC R20200;
DT 16-APR-1992 (first entry)
DE TGF-alpha-PE40Ab.
KW Pseudomonas exotoxin; bladder; mutant; target; receptor binding.
FH Key Location/Qualifiers
FT Region 5..54
FT /label= TGFalpha1-50
FT Region 59..420
FT /label= PE252-613
FT Misc_difference 176
FT /note= "Ser -> Thr"
FT Misc_difference 179
FT /note= "Cys -> Ala"
FT Misc_difference 186
FT /note= "Cys -> Ala"
PN EP-467536-A.
PD 22-JAN-1992.
PF 20-JUN-1991; 305582.
PR 21-JUN-1990; US-542281.
PR 14-MAR-1991; US-669269.
PA (MERI) MERCK & CO INC.

7. US-08-249-182-9 (1-16)

R20199 TGF-alpha-PE40aB.

ID R20199 standard; Protein; 419 AA.
 AC R20199;
 DT 16-APR-1992 (first entry)
 DE TGF-alpha-PE40aB.
 KW Pseudomonas exotoxin; bladder; mutant; target; receptor binding.
 FH Key Location/Qualifiers
 FT Region 5..54
 FT /label= TGFalpha1-50
 FT Region 58..419
 FT /label= PE252-613
 FT Misc_difference 71
 FT /note= "Cys -> Ala"
 FT Misc_difference 78
 FT /note= "Phe -> Leu"
 FT Misc_difference 93
 FT /note= "Cys -> Ala"
 PN EP-467536-A.
 PD 22-JAN-1992.
 PF 20-JUN-1991; 305582.
 PR 21-JUN-1990; US-542281.
 PR 14-MAR-1991; US-669269.
 PA (MERI) MERCK & CO INC.
 PI Ahern J, Heinbrook DC, Oliff AI, Stirdivant SM;
 DR WPI; 92-026359/04.
 PT Treatment of bladder cancer using hybrid protein - comprising
 PT cell targetting agent e.g. EGF that binds to EGF receptor on
 PT tumour cells and PE40 cell toxin
 PS Disclosure; Page 16-17; 34pp; English.
 CC The modified PE40 domains of the hybrid proteins have two or four of
 CC the Cys residues (designated Cys265, Cys287, Cys372 and Cys379)
 CC substituted with neutral amino acids, e.g. Gly, Ala, or Phe.
 CC TGF-alpha-PE40aB (R20199) has Cys265 and Cys287 replaced;
 CC TGF-alpha-PE40Ab (R20200) has Cys372 and Cys379 replaced; and
 CC TGF-alpha-PE40ab (R20201) has all four replaced.
 CC The modified hybrid proteins were produced in E.coli transformed
 CC with TAC expression vectors. Site specific mutations were introduced
 CC to the unmodified TGF-alpha-PE40 gene cloned in pTACTGF57-PE40.
 CC The mol. efficiently targets receptors on human bladder tumour cells
 CC (the modified PE40 domain has improved receptor binding) and is
 CC used for selectively killing bladder tumour cells.
 SQ Sequence 419 AA;
 SQ 56 A; 32 R; 9 N; 24 D; 0 B; 8 C; 21 Q; 32 E; 0 Z; 41 G; 11 H;
 SQ 13 I; 43 L; 4 K; 2 M; 12 F; 28 P; 22 S; 19 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||||| ||
CAGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNWTVERLLQAHRRQLEERGYV F V G Y H G T F L E A A Q S I
  190      200      210      220      230      X  240      250

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
  260      270      280      290      300

```

8. US-08-249-182-9 (1-16)

R07054 PE40AB protein comprising a portion of the Pseudon

6. US-08-249-182-9 (1-16)

R20201 TGF-alpha-PE40ab.

ID R20201 standard; Protein; 419 AA.
AC R20201;
DT 16-APR-1992 (first entry)
DE TGF-alpha-PE40ab.
KW Pseudomonas exotoxin; bladder; mutant; target; receptor binding.
FH Key Location/Qualifiers
FT Region 5..54
FT /label= TGFalpha1-50
FT Region 58..419
FT /label= PE252-613
FT Misc_difference 71
FT /note= "Cys -> Ala"
FT Misc_difference 78
FT /note= "Phe -> Leu"
FT Misc_difference 93
FT /note= "Cys -> Ala"
FT Misc_difference 175
FT /note= "Ser -> Thr"
FT Misc_difference 178
FT /note= "Cys -> Ala"
FT Misc_difference 185
FT /note= "Cys -> Ala"
PN EP-467536-A.
PD 22-JAN-1992.
PF 20-JUN-1991; 305582.
PR 21-JUN-1990; US-542281.
PR 14-MAR-1991; US-669269.
PA (MERI) MERCK & CO INC.
PI Ahern J, Heinbrook DC, Oliff AI, Stirdivant SM;
DR WPI; 92-026359/04.
PT Treatment of bladder cancer using hybrid protein - comprising
PT cell targetting agent e.g. EGF that binds to EGF receptor on
PT tumour cells and PE40 cell toxin
PS Disclosure; Page 20-21; 34pp; English.
CC The modified PE40 domains of the hybrid proteins have two or four of
CC the Cys residues (designated Cys265, Cys287, Cys372 and Cys372)
CC substituted with neutral amino acids, e.g. Gly, Ala, or Phe.
CC TGF-alpha-PE40aB (R20199) has Cys265 and Cys287 replaced;
CC TGF-alpha-PE40Ab (R20200) has Cys372 and Cys379 replaced; and
CC TGF-alpha-PE40ab (R20201) has all four replaced.
CC The modified hybrid proteins were produced in E.coli transformed
CC with TAC expression vectors. Site specific mutations were introduced
CC to the unmodified TGF-alpha-PE40 gene cloned in pTACTGF57-PE40.
CC The mol. efficiently targets receptors on human bladder tumour cells
CC (the modified PE40 domain has improved receptor binding) and is
CC used for selectively killing bladder tumour cells.
SQ Sequence 419 AA;
SQ 58 A; 32 R; 9 N; 24 D; 0 B; 6 C; 21 Q; 31 E; 0 Z; 42 G; 11 H;
SQ 13 I; 43 L; 4 K; 2 M; 12 F; 28 P; 21 S; 20 T; 5 W; 12 Y; 25 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

X 10 X
VNSMSTVFVGYGPTFK
||||| ||
AAGPADSGDALLERNYPTGAFLGDDGVSFSTRGTQNWTVRLLQAHRRGLEERGVFVGYHGTFLEAAQSI
190 200 210 220 230 X 240 250

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR

CC of cells for therapeutic, cytotoxic, diagnostic or other purposes.
CC See also R21436.7.
SQ Sequence 361 AA;
SQ 44 A; 30 R; 9 N; 19 D; 0 B; 4 C; 20 Q; 27 E; 0 Z; 38 G; 6 H;
SQ 13 I; 38 L; 5 K; 0 M; 11 F; 26 P; 18 S; 19 T; 6 W; 11 Y; 17 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||||| ||
CAGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNWTVERRLLQAHRLQLEERG Y V F V G Y H G T F L E A A Q S I
130      140      150      160      170      180      190 X

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
200      210      220      230      240
```

5. US-08-249-182-9 (1-16)

R04572 ORF4 product from the mos gene.

ID R04572 standard; protein; 407 AA.
AC R04572;
DT 14-SEP-1990 (first entry)
DE ORF4 product from the mos gene.
KW Rhizopine; mos gene; noc gene; nitrogen fixation; Medicago sativa.
OS Rhizobium meliloti strain L5-30.
PN AU8941262-A.
PD 15-MAR-1990.
PF 08-SEP-1988; A41262.
PR 08-SEP-1988; AU-000328.
PA (LUMI-) Luminis PTY Ltd.
PI Temp J, Kondorosi A, Putnoky P, Murphy PJ, Schell JS, De Bruijn FJ.
DR WPI; 90-139827/19.
DR N-PSDB; Q04303.
PT Bacteria contg. genes for rhizopine synthesis and catabolism - esp.
PT Rhizobium strains for increasing nitrogen fixation and growth in
PT leguminous plants.
PS Disclosure; p; English.
CC The mos ORF 4 product is a protein of a predicted size of 35.8kD.
CC Rhizobium strains, eg RM1021, containing the full mos gene and a noc gene
CC to catabolise rhizopine compounds are used to increase symbiotic nitrogen
CC fixation in Leguminaceae, esp. alfalfa. Where a noc gene is
CC present in separate bacteria both N-fixation and plant growth can be
CC promoted. Alternatively, mos genes are expressed in the plant and only
CC noc in the bacteria, this will cause desirable soil bacteria (eg being
CC used for biological control of a pathogen) to be held in the rhizosphere.
CC See also R04569-72.
SQ Sequence 407 AA;
SQ 46 A; 20 R; 11 N; 10 D; 0 B; 8 C; 6 Q; 11 E; 0 Z; 42 G; 11 H;
SQ 27 I; 49 L; 10 K; 17 M; 27 F; 16 P; 35 S; 21 T; 3 W; 6 Y; 31 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                || || | ||
GCRALIAISGSLASAGMTTALFMPSFAGKLAGFALLGLGMANLVPIIFSEASMNTVSKTVGLTFVSVCGYS
290      300      310      320      330      340      350 X

GFLVGPPPIIGASRRPLGSGELCSSSFAVGVIVACASVFFDRHRS
360      370      380      390      400
```

CC residues 365-378 of PE-40, in the N-terminus of domain Ib, which
 CC has ADP-ribosylation activity and is involved in protein synthesis.
 CC A chimeric gene, encoding the chimeric toxin, constructed in plasmid
 CC pWD154 was found to have similar ADP-ribosylation activity to the
 CC wild-type PE-40. Thus foreign amino-acid sequences can be
 CC introduced into the toxin while keeping intact the membrane
 CC translocation and cytotoxic activities. The implications are that
 CC chimeric proteins can be used to introduce foreign proteins such as
 CC enzymes, growth factors, lymphokines or antibodies into the cytosol
 CC of cells for therapeutic, cytotoxic, diagnostic or other purposes.
 CC See also R21435.7.

SQ Sequence 361 AA;
 SQ 43 A; 30 R; 9 N; 19 D; 0 B; 3 C; 19 Q; 27 E; 0 Z; 38 G; 6 H;
 SQ 13 I; 38 L; 5 K; 0 M; 11 F; 26 P; 21 S; 19 T; 6 W; 11 Y; 17 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                |||||  ||
CAGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNWTVRLLQAHRLGLEERGYV F V G Y H G T F L E A A Q S I
 130      140      150      160      170      180      190 X

VFGGVRRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
 200      210      220      230      240
  
```

4. US-08-249-182-9 (1-16)

R21435 PE-40 somatostatin substituted protein.

ID R21435 standard; protein; 361 AA.
 AC R21435;
 DT 16-JUN-1992 (first entry)
 DE PE-40 somatostatin substituted protein.
 KW Pseudomonas exotoxin-40; chimera; translocation; therapeutic;
 KW cytotoxic; diagnostic.
 OS Pseudomonas.
 FH Key Location/Qualifiers
 FT Modified_site 113..126
 FT /note= "modified somatostatin substitution,
 FT see comments"
 PN US7663455-A.
 PD 21-JAN-1992.
 PF 04-MAR-1991; 663455.
 PR 04-MAR-1991; US-663455.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Pastan IH;
 DR WPI; 92-079704/10.
 PT Recombinant chimeric proteins for diagnosis - contg. trans
 PT locating segment delivering foreign protein into cytosol of
 PT target cells
 PS Disclosure; Page 5; 45pp; English.
 CC The sequence shows residues 252 to 613 of Pseudomonas exotoxin 40,
 CC (retrieved from patent EP-383599-A) substituted at residue 113 by
 CC a synthetic somatostatin 14 molecule. This substitution is between
 CC residues 365-378 of PE-40, in the N-terminus of domain Ib, which
 CC has ADP-ribosylation activity and is involved in protein synthesis.
 CC A chimeric gene, encoding the chimeric toxin, constructed in plasmid
 CC pWD150 was found to have similar ADP-ribosylation activity to the
 CC wild-type PE-40. Thus foreign amino-acid sequences can be
 CC introduced into the toxin while keeping intact the membrane
 CC translocation and cytotoxic activities. The implications are that
 CC chimeric proteins can be used to introduce foreign proteins such as
 CC enzymes, growth factors, lymphokines or antibodies into the cytosol

PI Pastan IH;
 DR WPI; 92-079704/10.
 PT Recombinant chimeric proteins for diagnosis - contg. trans
 PT locating segment delivering foreign protein into cytosol of
 PT target cells
 PS Disclosure; Page 5; 45pp; English.
 CC The sequence shows residues 252 to 613 of Pseudomonas exotoxin 40,
 CC (retrieved from patent EP-383599-A) substituted at residue 113 by
 CC a synthetic methionine rich peptide. This substitution is between
 CC residues 365-380 of PE-40, in the N-terminus of domain 1b, which
 CC has ADP-ribosylation activity and is involved in protein synthesis.
 CC A chimeric gene, encoding the chimeric toxin, constructed in plasmid
 CC pWD163 was found to have similar ADP-ribosylation activity to the
 CC wild-type PE-40. Thus foreign amino-acid sequences can be
 CC introduced into the toxin while keeping intact the membrane
 CC translocation and cytotoxic activities. The implications are that
 CC chimeric proteins can be used to introduce foreign proteins such as
 CC enzymes, growth factors, lymphokines or antibodies into the cytosol
 CC of cells for therapeutic, cytotoxic, diagnostic or other purposes.
 CC See also R21435,7.
 SQ Sequence 361 AA;
 SQ 42 A; 30 R; 8 N; 20 D; 0 B; 4 C; 19 Q; 27 E; 0 Z; 38 G; 6 H;
 SQ 13 I; 38 L; 3 K; 9 M; 8 F; 27 P; 18 S; 18 T; 5 W; 11 Y; 17 V;

Initial Score = 7 Optimized Score = 7 Significance = 4.66
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                |||||  ||
CMGPADSGDALLERNYPTEAEFLGDGGDVSFSTRGTQNWTVERRLLQAHRRGLEERGVFVGYHGTFLEAAQSI
130      140      150      160      170      180      190 X

VFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQQDQEPDARGRIR
200      210      220      230      240
  
```

3. US-08-249-182-9 (1-16)

R21436 PE-40 somatostatin substituted protein.

ID R21436 standard; Protein; 361 AA.
 AC R21436;
 DT 16-JUN-1992 (first entry)
 DE PE-40 somatostatin substituted protein.
 KW Pseudomonas exotoxin-40; chimera; translocation; therapeutic;
 KW cytotoxic; diagnostic.
 OS Pseudomonas.
 FH Key Location/Qualifiers
 FT Modified_site 113..126
 FT /note= "modified somatostatin substitution,
 FT see comments"
 PN US7663455-A.
 PD 21-JAN-1992.
 PF 04-MAR-1991; 663455.
 PR 04-MAR-1991; US-663455.
 PA (USSH) US DEPT HEALTH & HUMAN.
 PI Pastan IH;
 DR WPI; 92-079704/10.
 PT Recombinant chimeric proteins for diagnosis - contg. trans
 PT locating segment delivering foreign protein into cytosol of
 PT target cells
 PS Disclosure; Page 5; 45pp; English.
 CC The sequence shows residues 252 to 613 of Pseudomonas exotoxin 40,
 CC (retrieved from patent EP-383599-A) substituted at residue 113 by
 CC a synthetic somatostatin 14 molecule. This substitution is between

1. US-08-249-182-9 (1-16)

R37451 Autotaxin peptide ATX 101.

ID R37451 standard; peptide; 16 AA.
AC R37451;
DT 22-JUL-1993 (first entry)
DE Autotaxin peptide ATX 101.
KW Cell motility stimulating; cancer metastasis; antibody; detection;
KW immunostains; disease outcome prediction; therapy choice;
KW cancer therapy; crosslinked toxins.
OS Synthetic.
PN US7822043-A.
PD 01-JAN-1993.
PF 17-JAN-1992; 822043.
PR 17-JAN-1992; US-822043.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Krutzsch H, Liotta LA, Schiffmann E, Stracke M.
DR WPI; 93-085861/10.
PT Motility stimulating protein named autotaxin - useful in cancer
PT diagnosis and therapy
PS Example; Page 33; 36pp; English.
CC The sequence is that of autotaxin peptide ATX 101. It may be used to
CC raise anti-autotaxin antibodies which can be used to diagnose cancer
CC metastasis and in immunostains of patient samples to detect the
CC presence of autotaxin. The level of autotaxin in tissue or body
CC fluids can be used to predict disease outcomes and/or choice of
CC therapy which may also include autotaxin inhibitors. Autotaxin
CC antibodies can be crosslinked to toxins (e.g. ricin A) for cancer
CC therapy.
SQ Sequence 16 AA;
SQ 0 A; 0 R; 1 N; 0 D; 0 B; 0 C; 1 Q; 0 E; 0 Z; 2 G; 1 H;
SQ 0 I; 0 L; 1 K; 0 M; 2 F; 1 P; 1 S; 2 T; 0 W; 1 Y; 3 V;

Initial Score = 15 Optimized Score = 15 Significance = 10.88
Residue Identity = 93% Matches = 15 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSMQTVFVGYGPTFK
||| |||||
VNSMQTVFVGYGPTFK
X      10      X

```

2. US-08-249-182-9 (1-16)

R21437 PE-40 protein contg. a methionine rich peptide.

ID R21437 standard; Protein; 361 AA.
AC R21437;
DT 16-JUN-1992 (first entry)
DE PE-40 protein contg. a methionine rich peptide.
KW Pseudomonas exotoxin-40; chimera; translocation; therapeutic;
KW cytotoxic; diagnostic.
OS Pseudomonas.
FH Key Location/Qualifiers
FT Modified_site 113..128
FT /note= "methionine rich peptide substitution,
FT see comments"
PN US7663455-A.
PD 21-JAN-1992.
PF 04-MAR-1991; 663455.
PR 04-MAR-1991; US-663455.
PA (USSH) US DEPT HEALTH & HUMAN.

Times:

CPU
00:00:25.98Total Elapsed
00:00:26.00

Number of residues: 5287517
 Number of sequences searched: 42145
 Number of scores above cutoff: 3752

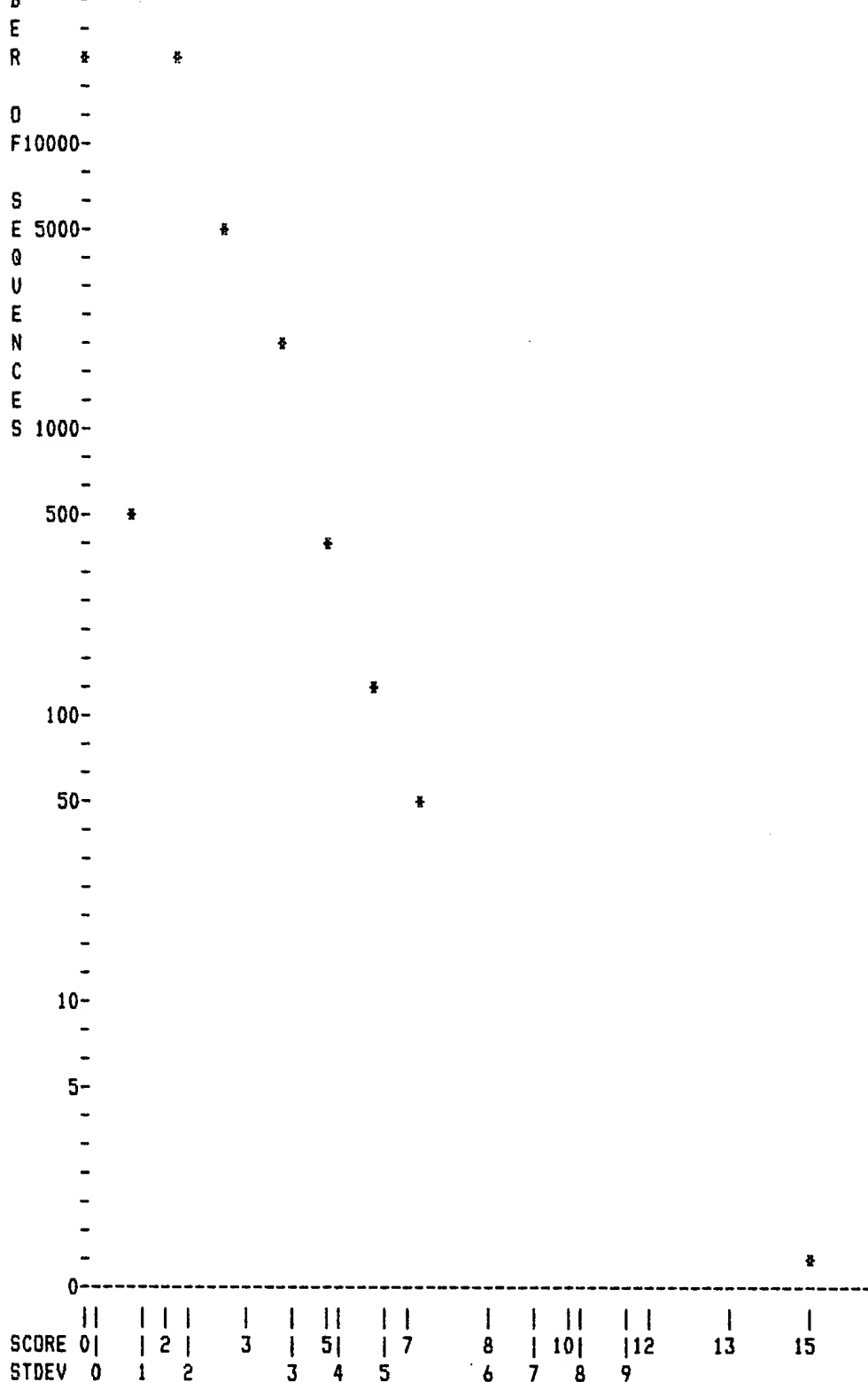
Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 10 standard deviations above mean ****						
1. R37451	Autotaxin peptide ATX 101.	16	15	15	10.88	0
**** 4 standard deviations above mean ****						
2. R21437	PE-40 protein contg. a methio	361	7	7	4.66	0
3. R21436	PE-40 somatostatin substitute	361	7	7	4.66	0
4. R21435	PE-40 somatostatin substitute	361	7	7	4.66	0
5. R04572	ORF4 product from the mos gen	407	7	7	4.66	0
6. R20201	TGF-alpha-PE40ab.	419	7	7	4.66	0
7. R20199	TGF-alpha-PE40aB.	419	7	7	4.66	0
8. R07054	PE40AB protein comprising a p	419	7	7	4.66	0
9. R20200	TGF-alpha-PE40Ab.	420	7	7	4.66	0
10. R06450	TGF-alpha-PE40-ab modified ps	420	7	7	4.66	0
11. R06449	TGF-alpha-PE40-Ab modified ps	420	7	7	4.66	0
12. R06448	TGF-alpha-PE40-aB modified ps	420	7	7	4.66	0
13. R06447	TGF-57-Pseudomonas exotoxin 4	420	7	7	4.66	0
14. R06994	PE40ab protein comprising a p	420	7	7	4.66	0
15. R06992	PE40aB protein comprising a p	420	7	7	4.66	0
16. R06993	PE40Ab protein comprising a p	420	7	7	4.66	0
17. R04934	Immunotoxin hybrid of human i	496	7	7	4.66	0
18. P70141	Sequence of a region of the S	537	7	7	4.66	0
19. R04920	Immunoprotein PEX46.	549	7	7	4.66	0
20. R04923	Immunoprotein TANG11.	557	7	7	4.66	0
21. P50501	Sequence of bovine parainflue	558	7	7	4.66	0
22. R31957	Sequence encoded by parainflu	572	7	7	4.66	0
23. R06023	Viral haemagglutinin neuramin	572	7	7	4.66	0
24. P94800	Parainfluenzae-3 gene product	572	7	7	4.66	0
25. P94799	Perdue strain of transmissabl	572	7	7	4.66	0
26. R04919	Immunoprotein PEX45.	574	7	7	4.66	0
27. R04924	Immunoprotein TANG12.	577	7	7	4.66	0
28. R40113	Pseudomonas exotoxin (S245C).	613	7	7	4.66	0
29. R40112	Pseudomonas exotoxin (K223C).	613	7	7	4.66	0
30. R40111	Pseudomonas exotoxin (S192C).	613	7	7	4.66	0
31. R40110	Pseudomonas exotoxin (S188C).	613	7	7	4.66	0
32. R40109	Pseudomonas exotoxin (R182C).	613	7	7	4.66	0
33. R40108	Pseudomonas exotoxin (S158C).	613	7	7	4.66	0
34. R40107	Pseudomonas exotoxin (S96C).	613	7	7	4.66	0
35. R40106	Pseudomonas exotoxin (S88C).	613	7	7	4.66	0
36. R40105	Pseudomonas exotoxin (S25C).	613	7	7	4.66	0
37. R40104	Pseudomonas exotoxin (K20C).	613	7	7	4.66	0
38. R40102	Pseudomonas exotoxin for site	613	7	7	4.66	0
39. R26983	(FRP51)-ETA fusion protein.	637	7	7	4.66	0



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

RA SHIMONAKA M.; SCHROEDER R.; SHIMASAKI S.; LING N.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 165:189-195(1989).
 CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
 CC PROMOTING EFFECTS OF THE IGFS ON CELLS CULTURE. THEY ALTER THE
 CC INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.
 CC -!- SUBUNIT: IGFBP-3 CAN CIRCULATE IN SERUM BOUND TO BOTH AN IGF
 CC PEPTIDE AND AN 80 KD GLYCOPROTEIN AS PART OF A 150 KD COMPLEX.
 CC -!- BINDS IGF-II MORE THAN IGF-I.
 CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
 DR EMBL; M31837; RNIGFBP3.
 DR EMBL; M33300; RNIGFB01.
 DR PIR; A30820; A30820.
 DR PIR; B30820; B30820.
 DR PIR; A26832; A26832.
 DR PIR; A34651; A34651.
 DR PIR; A33570; A33570.
 DR PIR; A36748; A36748.
 DR PROSITE; PS00222; IGF_BINDING.
 DR PROSITE; PS00484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 27
 FT CHAIN 28 292 INSULIN-LIKE GROWTH FACTOR BINDING
 FT PROTEIN 3.
 FT CARBOHYD 118 118 POTENTIAL.
 FT CARBOHYD 124 124 POTENTIAL.
 FT CARBOHYD 137 137 POTENTIAL.
 FT CARBOHYD 200 200 POTENTIAL.
 FT DOMAIN 237 286 THYROGLOBULIN TYPE I.
 FT CONFLICT 8 8 MISSING (IN REF. 2).
 SQ SEQUENCE 292 AA; 31680 MW; 437334 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 GGQPLWITATK
 |||| ||
 EDTLNHLKFLNVLSPRGVHIPNCDKKGFYKKKQCRPSKGRKRGFCWCVDKYGGQPLPGYDTKGKDDVHCLSVQ
 220 230 240 250 260 270 280 290

SQ

> 0 <
 0| |0 IntelliGenetics
 > 0 <

FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Seq. 9

Results file u249_9a.res made by on Thu 22 Sep 94 10:27:33-PDT.

Query sequence being compared: US-08-249-182-9 (1-16)
 Number of sequences searched: 42145
 Number of scores above cutoff: 3752

Results of the initial comparison of US-08-249-182-9 (1-16) with:
 Data bank : A-GeneSeq 15, all entries

100000-

N -

U50000-

M -

DA PROSITE; P500484; THYROGLOBULIN_1.
 KW GROWTH FACTOR BINDING; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 27
 FT CHAIN 28 291 INSULIN-LIKE GROWTH FACTOR BINDING
 FT PROTEIN 3.
 FT DOMAIN 28 134 IGF-BINDING (POTENTIAL).
 FT DOMAIN 138 161 SER/THR-RICH.
 FT DOMAIN 192 208 SER/THR-RICH.
 FT CARBOHYD 116 116 POTENTIAL.
 FT CARBOHYD 136 136 POTENTIAL.
 FT CARBOHYD 199 199 POTENTIAL.
 FT DOMAIN 236 285 THYROGLOBULIN TYPE I.
 SQ SEQUENCE 291 AA; 31660 MW; 441418 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 GGQPLWITATK
 ||| |
 EDTLNHLKFLNVLSPRGVHIPNCDKKGFYKKKQCRPSKGRKRGFCWCVDKYGGPLPGYTTKGKEDVHCYSM0
 220 230 240 250 260 270 280

SK
 290

15. US-08-249-182-8 (1-11)

IBP3_RAT INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECU

ID IBP3_RAT STANDARD; PRT; 292 AA.
 AC P15473;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR (IGFBP-3)
 DE (IBP-3) (IGF-BINDING PROTEIN 3).
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90088511
 RA SHIMASAKI S., KOBA A., MERCADO M., SHIMONAKA M., LING N.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 165:907-912(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RM 90147804
 RA ALBISTON A.L., HERINGTON A.C.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 166:892-897(1990).
 RN [3]
 RP SEQUENCE OF 28-68.
 RM 89050156
 RA ZAPF J., BORN W., CHANG J.Y., JAMES P., FROESCH E.R., FISCHER J.A.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 156:1187-1194(1988).
 RN [4]
 RP SEQUENCE OF 28-42.
 RM 87326380
 RA BAXTER R.C., MARTIN J.L.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 147:408-415(1987).
 RN [5]
 RP SEQUENCE OF 28-49.
 RC TISSUE=SERUM;
 RM 90073708

P1 CHAIN 80 283 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
SQ SEQUENCE 265 AA; 28561 MW; 365422 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG0PLWITATK
                                ||  |||
GAYSGKTDSEGGFESDAVAIANVLETSSAEVGGKPPYYLSVLTRTADGNEGKGHLITATVNDGKLYICKA0
170      180      190      200      210      220      230      240

AGDKRWFKGAKKFVENTATSFSLA
      250      260
```

14. US-08-249-182-8 (1-11)

IBP3_HUMAN INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECU

ID IBP3_HUMAN STANDARD; PRT; 291 AA.
AC P17936;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR (IGFBP-3)
DE (IBP-3) (IGF-BINDING PROTEIN 3).
GN IBP3.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RM 90324259
RA CUBBAGE M.L., SUWANICHKUL A., POWELL D.R.;
RL J. BIOL. CHEM. 265:12642-12649(1990).
RN [2]
RP SEQUENCE FROM N.A.
RM 89112197
RA WOOD W.I., CACHIANES G., HENZEL W.J., WINSLOW G.A., SPENCER S.A.,
RA HELLMISS R., MARTIN J.L., BAXTER R.C.;
RL MOL. ENDOCRINOL. 2:1176-1185(1988).
RN [3]
RP SEQUENCE OF 28-65.
RM 90368661
RA ZAPF J., KIEFER M., MERRYWEATHER J., MASIARZ F., BAUER D., BORN W.,
RA FISCHER J.A., FORESCH E.R.;
RL J. BIOL. CHEM. 265:14892-14898(1990).
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFS ON CELLS CULTURE. THEY ALTER THE
CC INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.
CC -!- SUBUNIT: IGFBP-3 CAN CIRCULATE IN SERUM BOUND TO BOTH AN IGF
CC PEPTIDE AND AN 80 KD GLYCOPROTEIN AS PART OF A 150 KD COMPLEX.
CC -!- TISSUE SPECIFICITY: EXPRESSED BY MOST TISSUES.
CC -!- DEVELOPMENTAL STAGE: IGFBP-3 LEVELS ARE HIGHER DURING EXTRAUTERINE
CC LIFE AND PEAK DURING PUBERTY.
CC -!- INDUCTION: IGFBP-3 LEVELS INCREASE IN THE PRESENCE OF IGF-I,
CC INSULIN AND OTHER GROWTH-STIMULATING FACTORS SUCH AS GROWTH
CC HORMONE, EPIDERMAL GROWTH FACTOR, AND PHORBOL ESTERS.
CC -!- BINDS IGF-II MORE THAN IGF-I.
CC -!- SIMILARITY: TO OTHER INSULIN-LIKE GROWTH FACTOR BINDING PROTEINS.
DR EMBL; M35878; HSIBP3.
DR PIR; A36578; IOHU3.
DR MIM; 146732; TENTH EDITION.
DR PROSITE; PS00222; IGF_BINDING.

CC WITH THE PHOTOSYSTEM II COMPLEX.
 CC -!- SIMILARITY: TO OTHER DEE2 SUBUNIT AND TO 16 KD SUBUNIT PROTEIN
 CC OF PHOTOSYSTEM II OF SPINACH.
 DR EMBL; X17213; SA23KDP2.
 DR EMBL; Y07498; SADES23.
 DR PIR; S03888; S03888.
 DR PIR; S10016; S10016.
 KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST; TRANSIT PEPTIDE;
 KW THYLAKOID MEMBRANE.
 FT TRANSIT 1 74 CHLOROPLAST.
 FT CHAIN 75 260 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
 SQ SEQUENCE 260 AA; 27925 MW; 347514 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                GG0PLWITATK
                                ||  |||
QAYFGETASEGGFDNNAVATANILETNIQDVGGKPPYYLSVLTRTADGDEGKHLITATVNGGKLYICKA9
  170      180      190      200      210  X  220  X  230

AGDKRWFKGANKFVEKAATSFSVA
  240      250      260

```

13. US-08-249-182-8 (1-11)

PSBP_TOBAC OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2)

ID PSBP_TOBAC STANDARD; PRT; 265 AA.
 AC P18212;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2) (23 KD SUBUNIT OF
 DE OXYGEN-EVOLVING SYSTEM OF PHOTOSYSTEM II).
 GN PSBP.
 OS NICOTIANA TABACUM (COMMON TOBACCO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR1; TISSUE=LEAF;
 RM 91329711
 RA HUA S., DUBE S.K., BARNETT N.M., KUNG S.;
 RL PLANT MOL. BIOL. 16:749-750(1991).
 RN [2]
 RP SEQUENCE OF 80-91.
 RC STRAIN=CV. KY57;
 RM 91329702
 RA TAKAHASHI H., EHARA Y., HIRANO H.;
 RL PLANT MOL. BIOL. 16:689-698(1991).
 CC -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 CC PHOTOSYSTEM II.
 CC -!- INDUCTION: BY LIGHT.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
 CC WITH THE PHOTOSYSTEM II COMPLEX.
 CC -!- SIMILARITY: TO OTHER DEE2 SUBUNIT AND TO 16 KD SUBUNIT PROTEIN
 CC OF PHOTOSYSTEM II OF SPINACH.
 DR EMBL; X55354; NTPH11.
 DR PIR; S14994; S14994.
 DR PIR; S15005; S15005.
 KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST; TRANSIT PEPTIDE;
 KW THYLAKOID MEMBRANE.
 FT TRANSIT 1 79 CHLOROPLAST.

RA MURATA N., KAJIURA H., FUJINURA Y., MIYAU M., MURATA T., WATANABE A.,
 RA SHINOZAKI K.;
 RL PRG. PHOTOSYNTH. RES. 1:701-704(1987).
 CC -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 CC PHOTOSYSTEM II.
 CC -!- INDUCTION: BY LIGHT.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
 CC WITH THE PHOTOSYSTEM II COMPLEX.
 CC -!- SIMILARITY: TO OTHER DEE2 SUBUNIT AND TO 16 KD SUBUNIT PROTEIN
 CC OF PHOTOSYSTEM II OF SPINACH.
 DR EMBL; X15552; PS23KDAP.
 DR EMBL; D13296; PSPSPBP.
 DR PIR; JS0771; JS0771.
 DR PIR; S03271; S03271.
 DR PIR; S07467; S07467.
 KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST; TRANSIT PEPTIDE;
 KW THYLAKOID MEMBRANE.
 FT TRANSIT 1 73 CHLOROPLAST.
 FT CHAIN 74 259 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
 FT CONFLICT 107 107 W -> K (IN REF. 2).
 SQ SEQUENCE 259 AA; 28047 MW; 348398 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 GGQPLWITATK
 || ||||
 QAFFGQTDSEGGFDNAVAVANILESSAPVIGGQYYNISVLTRTADGDEGGKHQLITATVKDGKLYICKAQ
 170 180 190 200 210 X 220 X 230
 AGDKRWFKCARKFVEDTASSFSVA
 240 250

12. US-08-249-182-8 (1-11)

PSBP_SINAL OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2)

ID PSBP_SINAL STANDARD; PRT; 260 AA.
 AC P11594;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-AUG-1991 (REL. 19, LAST ANNOTATION UPDATE)
 DE OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2) (23 KD SUBUNIT OF
 DE OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II).
 GN PSBP.
 OS SINAPIS ALBA (WHITE MUSTARD).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC CAPPARALES; CRUCIFERAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SEEDLING;
 RM 91346684
 RA MERKLE T., KRENZ M., WENNG A., SCHAEFER E.;
 RL PLANT MOL. BIOL. 14:889-890(1990).
 RN [2]
 RP SEQUENCE OF 13-260 FROM N.A.
 RC TISSUE=SEEDLING;
 RM 89211386
 RA WENNG A., EHMAN B., SCHAEFER E.;
 RL FEBS LETT. 246:140-144(1989).
 CC -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 CC PHOTOSYSTEM II.
 CC -!- INDUCTION: BY LIGHT.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED

DE OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II.
 GN PSBP OR PSBX.
 OS LYCOPERSICON ESCULENTUM (TOMATO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92256823
 RA BETTS S.D., PICHERSKY E.;
 RL PLANT MOL. BIOL. 18:995-996(1992).
 CC -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 CC PHOTOSYSTEM II.
 CC -!- INDUCTION: BY LIGHT.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
 CC WITH THE PHOTOSYSTEM II COMPLEX.
 CC -!- SIMILARITY: TO OTHER DEE2 SUBUNIT AND TO 16 KD SUBUNIT PROTEIN
 CC OF PHOTOSYSTEM II OF SPINACH.
 DR EMBL; X63007; LEPsBXMR.
 DR PIR; S20872; F2TOX2.
 KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST; TRANSIT PEPTIDE;
 KW THYLAKOID MEMBRANE.
 FT TRANSIT 1 73 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 74 258 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
 SQ SEQUENCE 258 AA; 27792 MW; 343591 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 GGQPLWITATK
 || ||||
 GAYFGKTDSEGGFESGAVATRNLLLEASSATVGGKEYYYLSVLTRTADGDEGGKHQLITATVNDGKLYICKAQ
 170 180 190 200 210 X 220 X 230
 AGDKRMFKGAKKFVENAATSFSIA
 240 250

11. US-08-249-182-8 (1-11)

PSBP_PEA OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2)

ID PSBP_PEA STANDARD; PRT; 259 AA.
 AC P16059;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (DEE2) (23 KD SUBUNIT OF
 DE OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II).
 GN PSBP.
 OS PISUM SATIVUM (GARDEN PEA).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
 OC FABACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. LITTLE MARVEL;
 RM 91370839
 RA WALES R., NEWMAN B.J., ROSE S.A., PAPPIN D., GRAY J.C.;
 RL PLANT MOL. BIOL. 13:573-582(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. ALASKA;
 RA KONISHI T., MARUTA Y., MURASE M., SHINOHARA K., WATANABE A.;
 RL SUBMITTED (OCT-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE OF 74-99.

HDGRQHDELRPITFDLDFISHPEGSVLITAGNTKVICNASVEDRVPPFLRGGGKGWITA EYSMLPRATNQRT

10 20 30 40 50 X 60 X 70

IRESSKGKISGRTEIQRILIGRALRAVVDLEKLGERTIW

80 90 100 110

9. US-08-249-182-8 (1-11)

PSBP_WHEAT OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (23 K

ID PSBP_WHEAT STANDARD; PRT; 258 AA.
AC Q00434;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (23 KD SUBUNIT OF
DE OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II).
GN PSBP.
OS TRITICUM AESTIVUM (WHEAT).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
OC CYPERALES; GRAMINEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVALON; TISSUE=LEAF;
RM 91329731
RA JAMES H.E., ROBINSON C.;
RL PLANT MOL. BIOL. 17:179-182(1991).
CC -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
CC PHOTOSYSTEM II.
CC -!- INDUCTION: BY LIGHT.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
CC WITH THE PHOTOSYSTEM II COMPLEX.
CC -!- SIMILARITY: TO OTHER DEE2 SUBUNIT AND TO 16 KD SUBUNIT PROTEIN
CC OF PHOTOSYSTEM II OF SPINACH.
DR EMBL; X57407; TAPPSBP.
DR PIR; S22763; S22763.
KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST; TRANSIT PEPTIDE;
KW THYLAKOID MEMBRANE.
FT TRANSIT 1 73 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 74 258 OXYGEN-EVOLVING ENHANCER PROTEIN 2.
SQ SEQUENCE 258 AA; 27269 MW; 341934 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10
GGQPLWITATK
|| |||

QSYGCKTDSEGGFESDAVATANVLESSAPVVDGKQYYSITVLTTRTADGDEGGKHQLITATVADGKLYVCKAQ
170 180 190 200 210 X 220 X 230

RDKRWFKGAKKFVENAAGSFSVA
240 250

10. US-08-249-182-8 (1-11)

PSBP_LYCES OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (23 K

ID PSBP_LYCES STANDARD; PRT; 258 AA.
AC P29795;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE OXYGEN-EVOLVING ENHANCER PROTEIN 2 PRECURSOR (23 KD SUBUNIT OF

CC SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC -!- CAUTION: BE CAREFUL OF POSSIBLE CONFUSION BETWEEN PRC-III AND
 CC PLP-III, THE TWO PROTEINS, WHILE DIFFERENT, ARE CALLED PLACENTAL
 CC PROLACTIN-RELATED PROTEIN III.
 DR EMBL; M27240; M27240.
 DR PIR; B34078; B34078.
 DR PROSITE; PS00266; SOMATOTROPIN_1.
 DR PROSITE; PS00338; SOMATOTROPIN_2.
 KW HORMONE; PLACENTA; SIGNAL.
 FT SIGNAL 1 13 POTENTIAL.
 FT CHAIN 14 213 PLACENTAL PROLACTIN-RELATED PROTEIN III.
 FT DISULFID 72 190 BY SIMILARITY.
 FT DISULFID 207 213 BY SIMILARITY.
 SQ SEQUENCE 213 AA; 25016 MW; 236080 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X X
 GG0PLWITATK
 || | ||
 QVNSCPSCCPDVFDIPLES LHLFLNASRLSHDIVNHTTINFHEFDEKYAQNQPYTINATKSCHTNSLHTPQ
 10 20 30 40 50 60 70 80
 EREKALRMNEDLSKWILMLLYSWHRPLYLLVKDLQSMK
 90 100 110 120

8. US-08-249-182-8 (1-11)

RNPH_BACSU RIBONUCLEASE PH (EC 2.7.7.56) (RNASE PH) (TRNA

ID RNPH_BACSU STANDARD; PRT; 245 AA.
 AC P28619;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE RIBONUCLEASE PH (EC 2.7.7.56) (RNASE PH) (TRNA
 DE NUCLEOTIDYLTRANSFERASE).
 GN RPH.
 OS BACILLUS SUBTILIS.
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92325065
 RA CRAVEN M.G., HENNER D.J., ALESSI D., SCHAUER A.T., OST K.A.,
 RA DEUTSCHER M.P., FRIEDMAN D.I.;
 RL J. BACTERIOL. 174:4727-4735(1992).
 CC -!- FUNCTION: RNASE PH IS A PHOSPHOROLYTIC EXORIBONUCLEASE THAT
 CC REMOVES NUCLEOTIDE RESIDUES FOLLOWING THE -CCA TERMINUS OF TRNA
 CC AND ADDS NUCLEOTIDES TO THE ENDS OF RNA MOLECULES BY USING
 CC NUCLEOSIDE DIPHOSPHATES AS SUBSTRATES.
 CC -!- CATALYTIC ACTIVITY: TRNA(N+1) + ORTHOPHOSPHATE = TRNA(N) +
 CC A NUCLEOSIDE DIPHOSPHATE.
 CC -!- SIMILARITY: TO OTHER SPECIES RNASES PH.
 DR EMBL; M85163; BSRPHA.
 DR PIR; A44914; A44914.
 KW TRANSFERASE; NUCLEOTIDYLTRANSFERASE; TRNA PROCESSING.
 SQ SEQUENCE 245 AA; 26681 MW; 287533 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 GG0PLWITATK

ID SSBP_ECOLI STANDARD; PRT: 174 AA.
AC P18022;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABILIZING PROTEIN).
GN SSB.
OS ESCHERICHIA COLI.
OG PLASMID COLIB-P9.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 89213928
RA HOWLAND C.J., REES C.E.D., BARTH P.T., WILKINS B.M.;
RL J. BACTERIOL. 171:2466-2473(1989).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE CONJUGATIVE PROCESSING OF DNA. IT
CC HAS A FUNCTIONAL RELATIONSHIP WITH PSI (PLASMID-MEDIATED SOS
CC INHIBITION) PROTEINS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHERS PROKARYOTIC AND MITOCHONDRIAL SS-DNA
CC BINDING PROTEINS.
DR EMBL; M25505; M25505.
DR PIR; A32304; DDECIB.
DR PROSITE; PS00735; SSB_1.
DR PROSITE; PS00736; SSB_2.
KW PLASMID; DNA-BINDING; DNA REPLICATION.
FT INIT_MET 0 0 BY SIMILARITY.
FT DNA_BIND 54 60 BY SIMILARITY.
SQ SEQUENCE 174 AA; 19109 MW; 144724 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GGQPLWITATK
                                |||    |||
YIEGQLRTRSWDDNGITRYITEILVKTTGTNQLGSAPOQNAQAQPKPOQNGQPOSADATKKGGAKTKGRGR
 80      90      100      110      120      130      140

KAAQPEPQPTPEGEDYGFSDDIPF
150      160      170

```

7. US-08-249-182-8 (1-11)

PRR3_BOVIN PLACENTAL PROLACTIN-RELATED PROTEIN III PRECURSOR

ID PRR3_BOVIN STANDARD; PRT: 213 AA.
AC P12402;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE PLACENTAL PROLACTIN-RELATED PROTEIN III PRECURSOR (PRC-III).
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RM 89352599
RA KESSLER M.A., MILOSAVLJEVIC M., ZIELER C.G., SCHULER L.A.;
RL BIOCHEMISTRY 28:5154-5161(1989).
CC -!- FUNCTION: PLACENTAL PROLACTIN-RELATED PROTEINS MAY PLAY A
CC SPECIFIC ROLE DURING GESTATION.
CC -!- SUBCELLULAR LOCATION: SECRETED.

FL CONFLICT 77 77 S -7 A (IN REF. 3).
SQ SEQUENCE 170 AA; 17461 MW; 168534 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                ||| | ||
PASKTFESYRVMTGVHTNDATKKVIVKLADTPQLTDVLNSTV@MPISVSWGG@VLSTTAKEFEAAALGYSAS
 60      70      80      90      100      110      120
```

```

GVNGVSSSGELVISAAPKTAGTAPTAGNYSGVVSLVMTL
130      140      150      160
```

5. US-08-249-182-8 (1-11)

SSBR_ECOLI SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABI

ID SSBR_ECOLI STANDARD; PRT; 174 AA.
AC P28045;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABILIZING PROTEIN).
GN SSB.
OS ESCHERICHIA COLI.
OG PLASMID R64.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 91180109
RA RUVOLO P.P., KEATING K.M., WILLIAMS K.R., CHASE J.W.;
RL PROTEINS 9:120-134(1991).
CC -!- FUNCTION: MAY CONTRIBUTE TO THE CONJUGATIVE PROCESSING OF DNA. IT
CC HAS A FUNCTIONAL RELATIONSHIP WITH PSI (PLASMID-MEDIATED SOS
CC INHIBITION) PROTEINS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHERS PROKARYOTIC AND MITOCHONDRIAL SS-DNA
CC BINDING PROTEINS.
DR PIR; A38487; A38487.
DR PROSITE; PS00735; SSB_1.
DR PROSITE; PS00736; SSB_2.
KW DNA-BINDING; DNA REPAIR; DNA REPLICATION.
FT INIT_MET 0 0 BY SIMILARITY.
FT DNA_BIND 54 60 BY SIMILARITY.
SQ SEQUENCE 174 AA; 19181 MW; 144576 CN;

Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      X
                                GG@PLWITATK
                                ||| |||
YIEG@LRTRSWDDNGITRYITEILVKTTGT@MLGSAP@@NAGA@PK@@NG@@SADATKKGGAKTKGRER
 80      90      100      110      120      130      140
```

```

KAA@PEP@P@TPEGEDYGFSDDIPF
150      160      170
```

6. US-08-249-182-8 (1-11)

SSBP_ECOLI SINGLE-STRAND BINDING PROTEIN (SSB) (HELIX-DESTABI

SEQUENCE 125 AA; 14535 RW; 92013 CN;
Initial Score = 6 Optimized Score = 6 Significance = 4.52
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

          X      10
          GGQPLWITATK
          || ||  ||
PRTEKLSKFLDDDTFGKFQEVVGYNPLQVLRYDLGGKPLLYAETKVDILSTVPRPGYNPSSQRIFEMQLMPK
      10      20      30  X  40  X  50      60      70

MIFDLEEVVSVDNMGMEWGTLVF
      80      90

```

4. US-08-249-182-8 (1-11)

FMC1_ECOLI CFA/I FIMBRIAL SUBUNIT B PRECURSOR (COLONISATION F

ID FMC1_ECOLI STANDARD; PRT; 170 AA.
AC P02971;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE CFA/I FIMBRIAL SUBUNIT B PRECURSOR (COLONISATION FACTOR ANTIGEN I
DE SUBUNIT B) (CFA/I PILIN) (CFA/I ANTIGEN).
GN CFAB.
OS ESCHERICHIA COLI.
OG PLASMID NTP513.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 89173309
RA KARJALAINEN T.K., EVANS D.G., SO M., LEE C.H.;
RL INFECT. IMMUN. 57:1126-1130(1989).
RN [2]
RP SEQUENCE FROM N.A.
RM 89330163
RA HAMERS A.M., PEL H.J., WILLSHAW G.A., KUSTERS J.G.,
RA VAN DER ZEIJST B.A.M., GAASTRA W.;
RL MICROB. PATHOG. 6:297-309(1989).
RN [3]
RP SEQUENCE OF 24-170 FROM N.A.
RC STRAIN=H-10407;
RM 82235736
RA KLEMM P.;
RL EUR. J. BIOCHEM. 124:339-348(1982).
CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC -!- SUBUNIT: CFA/I FIMBRIAE ARE RATHER RIGID, THREAD-LIKE FILAMENTS OF
CC 0.5-1 MICROMETER, WITH AN APPARENT AXIAL HOLE, AND A DIAMETER OF
CC 7 NANOMETERS. A SINGLE CFA/I FIMBRIA CONSISTS OF ABOUT 100
CC IDENTICAL PROTEIN SUBUNITS.
CC -!- INDUCTION: CFA/I FIMBRIAE ARE ONLY EXPRESSED IN THE PRESENCE OF
CC THE POSITIVE REGULATOR CFAD.
CC -!- SIMILARITY: TO THE CS1 FIMBRIAL SUBUNIT A (CSOA).
DR EMBL; M55661; ECCFAIA.
DR PIR; A30589; YQECCL.
KW FIMBRIA; ANTIGEN; PLASMID; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 24 170 CFA/I FIMBRIAL PROTEIN B.
FT CONFLICT 37 37 V -> A (IN REF. 2).
FT CONFLICT 76 76 D -> N (IN REF. 3).

CC -!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + NADP(+) = 2,3-
 CC DEHYDROACYL-[ACYL-CARRIER PROTEIN] + NADPH.
 CC -!- CATALYTIC ACTIVITY: OLEOYL-[ACYL-CARRIER PROTEIN] + H(2)O =
 CC ACYL-CARRIER PROTEIN + OLEATE.
 CC -!- SUBUNIT: HOMODIMER, WHICH ARE ARRANGED IN A HEAD TO TAIL FASHION.
 DR EMBL; J03860; GGFAS.
 DR EMBL; J04485; GGFASB.
 DR EMBL; J02839; GGFASA.
 DR PIR; A33918; XYCHFA.
 DR PIR; A32015; A32015.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE.
 KW FATTY ACID BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME; PHOSPHOPANTETHEINE;
 KW TRANSFERASE; HYDROLASE; OXIDOREDUCTASE; LIGASE; NADP.
 FT DOMAIN 1 338 BETA-KETOACYL SYNTHASE.
 FT DOMAIN ? ? ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN ? 1793 ENOYL REDUCTASE.
 FT DOMAIN 1794 2046 BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2047 2135 ACYL CARRIER.
 FT DOMAIN 2136 2446 THIOESTERASE.
 FT ACT_SITE 87 87 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 506 506 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 2084 2084 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT_SITE 2235 2235 THIOESTERASE.
 FT VARSPLIC 2276 2283 MISSING (IN ISOFORM 2).
 FT CONFLICT 1504 1504 R -> W (IN REF. 2).
 FT CONFLICT 1659 1659 E -> Q (IN REF. 2).
 FT CONFLICT 1672 1672 N -> S (IN REF. 2).
 SQ SEQUENCE 2446 AA; 267246 MW; 21121641 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.65
 Residue Identity = 63% Matches = 7 Mismatches = 4
 Gaps = 0 Conservative Substitutions = 0

X X
 GGQPLWITATK
 |||| |||

LNLVAMKRSFFGSVIFLCRRQSPAKAPILLPVDDTHYKWVDSLKEILADSSEQPLWLTATNCGNSGILGMVN
 1320 1330 1340 1350 1360 1370 1380

CLRLEAEGHRIRC VFVSNLSPSSTVPATSLSSLEMQKII
 1390 1400 1410 1420

3. US-08-249-182-8 (1-11)

YIFM_YEAST HYPOTHETICAL PROTEIN IN IFM1 3'REGION (FRAGMENT).

ID YIFM_YEAST STANDARD; PRT; 125 AA.
 AC P25040;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN IN IFM1 3'REGION (FRAGMENT).
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 92037620
 RA VAMBUTAS A., ACKERMAN S.H., TZAGOLOFF A.;
 RL EUR. J. BIOCHEM. 201:643-652(1991).
 DR EMBL; X58379; SCIFM1.
 DR PIR; S20178; S20178.
 DR PIR; S17024; S17024.
 KW HYPOTHETICAL PROTEIN.
 FT NON_TER 1 1

ID FAS_CHICK STANDARD; PRT: 2446 AA.
 AC P12276;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE FATTY ACID SYNTHASE (EC 2.3.1.85) (CONTAINS: EC 2.3.1.38, EC 2.3.1.39,
 DE EC 2.3.1.41, EC 1.1.1.100, EC 4.2.1.61, EC 1.3.1.10, AND EC 3.1.2.14).
 GN FAS.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE OF 1-1701 FROM N.A.
 RC TISSUE=LIVER;
 RM 89282777
 RA HOLZER K.P., LIU W., HAMMES G.G.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:4387-4391(1989).
 RN [2]
 RP SEQUENCE OF 1494-2438 FROM N.A.
 RM 89139426
 RA CHIRALA S.S., KASTURI R., PAZIRANDEH M., STOLOW D.T., HUANG W.-Y.,
 RA WAKIL S.J.;
 RL J. BIOL. CHEM. 264:3750-3757(1989).
 RN [3]
 RP SEQUENCE OF 1678-2438 FROM N.A.
 RM 88320436
 RA YUAN Z., LIU W., HAMMES G.G.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 85:6328-6331(1988).
 RN [4]
 RP SEQUENCE OF 2128-2438 FROM N.A.
 RM 89088152
 RA KASTURI R., CHIRALA S., PAZIRANDEH M., WAKIL S.J.;
 RL BIOCHEMISTRY 27:7778-7785(1988).
 RN [5]
 RP SEQUENCE OF 2047-2135.
 RM 89192401
 RA HUANG W.-Y., STOOPS J.K., WAKIL S.J.;
 RL ARCH. BIOCHEM. BIOPHYS. 270:92-98(1989).
 RN [6]
 RP SEQUENCE OF 2135-2434.
 RC STRAIN=WHITE LEGHORN;
 RM 89088151
 RA YANG C.-Y., HUANG W.-Y., CHIRALA S., WAKIL S.J.;
 RL BIOCHEMISTRY 27:7773-7777(1988).
 RN [7]
 RP SEQUENCE OF 593-600 AND 1624-1635.
 RM 89323081
 RA CHANG S.I., HAMMES G.G.;
 RL BIOCHEMISTRY 28:3781-3788(1989).
 CC -!- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -!- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH =
 CC LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+).
 CC -!- CATALYTIC ACTIVITY: ACETYL-COA + [ACYL-CARRIER PROTEIN] = COA
 CC + ACETYL-[ACYL-CARRIER PROTEIN].
 CC -!- CATALYTIC ACTIVITY: MALONYL-COA + [ACYL-CARRIER PROTEIN] = COA
 CC + MALONYL-[ACYL-CARRIER PROTEIN].
 CC -!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-
 CC CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +
 CC [ACYL-CARRIER PROTEIN].
 CC -!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
 CC NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.
 CC -!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYPALMITOYL-[ACYL-CARRIER PROTEIN]

16. VMAT_RINDR	MATRIX PROTEIN.	333	6	6	4.52	0
17. GPDA_SCHPD	GLYCEROL-3-PHOSPHATE DEHYDROG	384	6	6	4.52	0
18. PEPC_HUMAN	PROGASTRICSIN PRECURSOR (EC 3	388	6	6	4.52	0
19. PEPC_RAT	PROGASTRICSIN PRECURSOR (EC 3	392	6	6	4.52	0
20. AAT_BACSP	ASPARTATE AMINOTRANSFERASE (E	392	6	6	4.52	0
21. SBCE_ECOLI	EXONUCLEASE SBCE.	400	6	7	4.52	0
22. B3AR_RAT	BETA-3-ADRENERGIC RECEPTOR.	400	6	6	4.52	0
23. PPB3_BACSU	ALKALINE PHOSPHATASE III PREC	462	6	6	4.52	0
24. CYAA_TRYEQ	ADENYLATE CYCLASE (EC 4.6.1.1	469	6	6	4.52	0
25. ATPB_EUGGR	ATP SYNTHASE BETA CHAIN (EC 3	480	6	6	4.52	0
26. ATPB_CHLEL	ATP SYNTHASE BETA CHAIN (EC 3	481	6	6	4.52	0
27. NIFB_RHIME	NIFB PROTEIN.	490	6	6	4.52	0
28. ATPB_CHLRE	ATP SYNTHASE BETA CHAIN (EC 3	491	6	6	4.52	0
29. OM6E_CHLTR	60 KD OUTER MEMBRANE PROTEIN	547	6	6	4.52	0
30. ATP2_ORYSA	ATP SYNTHASE BETA CHAIN, MITO	551	6	6	4.52	0
31. ATP2_MAIZE	ATP SYNTHASE BETA CHAIN, MITO	553	6	6	4.52	0
32. ATP2_NICPL	ATP SYNTHASE BETA CHAIN, MITO	560	6	6	4.52	0
33. LCFA_ECOLI	LONG-CHAIN-FATTY-ACID--COA LI	561	6	6	4.52	0
34. ATP2_HEVBR	ATP SYNTHASE BETA CHAIN, MITO	562	6	6	4.52	0
35. MASY_CUCSA	MALATE SYNTHASE, GLYOXYSOMAL	568	6	6	4.52	0
36. YLP5_CAEEL	HYPOTHETICAL 66.5 KD PROTEIN	608	6	6	4.52	0
37. HEX3_YEAST	HEXOSE METABOLISM-RELATED PRO	619	6	6	4.52	0
38. PC1_MOUSE	PLASMA-CELL MEMBRANE GLYCOPRO	871	6	6	4.52	0
39. SCD5_YEAST	SCD5 PROTEIN.	872	6	7	4.52	0
40. CARB_BACSU	CARBAMOYL-PHOSPHATE SYNTHASE,	1071	6	6	4.52	0

1. US-08-249-182-8 (1-11)

VG02_VZVD HYPOTHETICAL GENE 2 PROTEIN.

ID VG02_VZVD STANDARD; PRT; 238 AA.
AC P09267;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL GENE 2 PROTEIN.
GN 2.
OS VARICELLA-ZOSTER VIRUS (STRAIN DUMAS) (VZV).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RM 86306657
RA DAVISON A.J., SCOTT J.E.;
RL J. GEN. VIROL. 67:1759-1816(1986).
CC -!- SIMILARITY: TO EHV-1 3.
DR EMBL; X04370; HEVZVXX.
DR PIR; B27212; WZBE2.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 238 AA; 25984 MW; 292124 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.65
Residue Identity = 63% Matches = 7 Mismatches = 4
Gaps = 0 Conservative Substitutions = 0

X 10
GGQPLWITATK
|| || || ||
SETLAYGHVPAFINGSTLVRPSLNATAEENPASETRCLLRVLAGRTVDLPGGGTLHITCTKYVVIICKYSKP
10 20 30 40 50 X 60 X 70
GERLSLARLIGRANTPGGARTFIILAKKEKRSTTLGYEC
80 90 100 110

2. US-08-249-182-8 (1-11)

FAS_CHICK FATTY ACID SYNTHASE (EC 2.3.1.85) (CONTAINS: EC 2.

	0	1	2	3	4	5	6	7
SCORE	0	1	2	3	4	5	6	7
STDEV	-1	0	1	2	3	5		

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.89
Times:	CPU	Total Elapsed	
	00:00:50.96	00:00:58.00	
Number of residues:	12496420		
Number of sequences searched:	36000		
Number of scores above cutoff:	3932		

Cut-off raised to 2.
Cut-off raised to 3.
Cut-off raised to 4.
Cut-off raised to 5.

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
1. VG02_VZVD	HYPOTHETICAL GENE 2 PROTEIN.	238	7	7	5.65	0
2. FAS_CHICK	FATTY ACID SYNTHASE (EC 2.3.1	2446	7	7	5.65	0
**** 4 standard deviations above mean ****						
3. YIFM_YEAST	HYPOTHETICAL PROTEIN IN IFM1	125	6	6	4.52	0
4. FMC1_ECOLI	CFA/I FIMBRIAL SUBUNIT B PREC	170	6	6	4.52	0
5. SSBP_ECOLI	SINGLE-STRAND BINDING PROTEIN	174	6	6	4.52	0
6. SSBP_ECOLI	SINGLE-STRAND BINDING PROTEIN	174	6	6	4.52	0
7. PRR3_BOVIN	PLACENTAL PROLACTIN-RELATED P	213	6	6	4.52	0
8. RNPH_BACSU	RIBONUCLEASE PH (EC 2.7.7.56)	245	6	6	4.52	0
9. PSBP_WHEAT	OXYGEN-EVOLVING ENHANCER PROT	258	6	6	4.52	0
10. PSBP_LYCES	OXYGEN-EVOLVING ENHANCER PROT	258	6	6	4.52	0
11. PSBP_PEA	OXYGEN-EVOLVING ENHANCER PROT	259	6	6	4.52	0
12. PSBP_SINAL	OXYGEN-EVOLVING ENHANCER PROT	260	6	6	4.52	0
13. PSBP_TOBAC	OXYGEN-EVOLVING ENHANCER PROT	265	6	6	4.52	0
14. IBP3_HUMAN	INSULIN-LIKE GROWTH FACTOR BI	291	6	6	4.52	0
15. IBP3_RAT	INSULIN-LIKE GROWTH FACTOR BI	292	6	6	4.52	0

240 250
> 0 <
0| 10 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_8s.res made by on Thu 22 Sep 94 10:19:13-PDT.

Query sequence being compared:US-08-249-182-8 (1-11)
Number of sequences searched: 36000
Number of scores above cutoff: 3932

Results of the initial comparison of US-08-249-182-8 (1-11) with:
Data bank : Swiss-Prot 28, all entries

100000-
-
N -
U50000-
M -
B -
E -
R -
-
O -
F10000-
-
S -
E 5000-
Q -
U -
E -
N *
C -
E -
S 1000-
-
-
500-
-
-
-
-
-
100-
-
-
50-
-
-
-
-
10-
-
5-
-
-
-

*

*

*

*

*

*

*

KEYWORDS chloroplast; photosynthesis; photosystem II; thylakoid

FEATURE

1-73 #domain transit peptide (chloroplast) #status predicted
#label TNP\

74-258 #protein photosystem II oxygen-evolving complex protein
2 #status predicted #label MAT

SUMMARY #length 258 #molecular-weight 27792 #checksum 9784

SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10
GGQPLWITATK
|| |||
GAYFGKTDSEGGFESGAVATRNLL EASSATVGGKEYVYLSVLTRTADGDEGGKHQLITATVNDGKLYICKAQ
170 180 190 200 210 X 220 X 230

AGDKRWFKGAKKFVENAATSFSIA
240 250

15. US-08-249-182-8 (1-11)

JS0771 photosystem II oxygen-evolving complex protein 2 p

ENTRY JS0771 #type complete

TITLE photosystem II oxygen-evolving complex protein 2 precursor -
garden pea

ALTERNATE_NAMES photosystem II 23K protein; photosystem II psbP protein

ORGANISM #formal_name Pisum sativum #common_name garden pea

DATE 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change
30-Sep-1993

ACCESSIONS JS0771

REFERENCE JS0771

#authors Konishi, T.; Maruta, Y.; Murase, M.; Shinohara, K.; Watanabe,
A.

#submission submitted to JIPID, October 1992

#contents Strain Alaska

#accession JS0771

##molecule_type mRNA

##residues 1-259 ##label KON

##cross-references DDBJ:D13296

##note the translation of the nucleotide sequence is not given
in this paper

GENETICS

#gene psbP

CLASSIFICATION #superfamily photosystem II oxygen-evolving complex protein 2

KEYWORDS chloroplast; photosynthesis; photosystem II; thylakoid

FEATURE

1-73 #domain transit peptide (chloroplast) #label TNP\

74-259 #protein photosystem II oxygen-evolving complex protein
2 #status experimental #label MAT

SUMMARY #length 259 #molecular-weight 28047 #checksum 3807

SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

X 10
GGQPLWITATK
|| |||
GAFFGQTDSEGGFDNAVAVANILESSAPVIGGQYNNISVLTRTADGDEGGKHQLITATVKDGKLYICKAQ
170 180 190 200 210 X 220 X 230

ENTRY S22763 #type complete
 TITLE photosystem II oxygen-evolving complex protein 2 precursor - wheat
 ALTERNATE_NAMES oxygen-evolving complex 23K protein
 ORGANISM #formal_name Triticum aestivum #common_name common wheat
 DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 03-Feb-1994
 ACCESSIONS S22763
 REFERENCE S22763
 #authors James, H.E.; Robinson, C.
 #journal Plant Mol. Biol. (1991) 17:179-182
 #title Nucleotide sequence of cDNA encoding the precursor of the 23 kDa protein of the photosynthetic oxygen-evolving complex from wheat.
 #cross-references MUID:91329731
 #accession S22763
 ##molecule_type mRNA
 ##residues 1-258 ##label JAM
 ##cross-references EMBL:X57407
 CLASSIFICATION #superfamily photosystem II oxygen-evolving complex protein 2
 KEYWORDS chloroplast; photosynthesis; photosystem II; thylakoid
 FEATURE
 1-73 #domain transit peptide (chloroplast) #status predicted #label TNP\
 74-258 #protein photosystem II oxygen-evolving complex protein 2 #status predicted #label MAT
 SUMMARY #length 258 #molecular-weight 27269 #checksum 5649
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                GGQPLWITATK
                                ||  |||
QSYGGKTDSEGGFESDAVATANVLESSAPVVDGKQYYSITVLTRTADGDEGGKHQLITATVADGKLYVCKAQ
  170      180      190      200      210  X  220  X  230

RDKRWFKGAKKFVENAAGSFSVA
  240      250

```

14. US-08-249-182-8 (1-11)

F2TOX2 photosystem II oxygen-evolving complex protein 2 p

ENTRY F2TOX2 #type complete
 TITLE photosystem II oxygen-evolving complex protein 2 precursor - tomato
 ALTERNATE_NAMES photosystem II oxygen-evolving complex 23K protein
 ORGANISM #formal_name Lycopersicon esculentum #common_name tomato
 DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 30-Jun-1993
 ACCESSIONS S20872
 REFERENCE S20872
 #authors Betts, S.; Pichersky, E.
 #journal Plant Mol. Biol. (1992) 18:995-996
 #title Nucleotide sequence of cDNA encoding the precursor of the 23 kDa photosystem II protein of tomato.
 #cross-references MUID:92256823
 #accession S20872
 ##molecule_type mRNA
 ##residues 1-258 ##label BET
 ##cross-references EMBL:X63007
 CLASSIFICATION #superfamily photosystem II oxygen-evolving complex protein 2

10. US-08-249-182-8 (1-11)
 C48652 transfer protein spdA - Streptomyces ambofaciens p

ENTRY C48652 #type complete
 TITLE transfer protein spdA - Streptomyces ambofaciens plasmid pSAM2
 ORGANISM #formal_name Streptomyces ambofaciens
 DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-May-1994
 ACCESSIONS C48652
 REFERENCE A48652
 #authors Hagege, J.; Pernodet, J.L.; Sezonov, G.; Gerbaud, C.; Friedmann, A.; Guerineau, M.
 #journal J. Bacteriol. (1993) 175:5529-5538
 #title Transfer functions of the conjugative integrating element pSAM2 from Streptomyces ambofaciens: characterization of a kil-kor system associated with transfer.
 #accession C48652
 ##status preliminary
 ##molecule_type DNA
 ##residues 1-224 ##label HAG
 ##cross-references ENBL:Z19593

GENETICS
 #genome plasmid
 SUMMARY #length 224 #molecular-weight 23575 #checksum 9089
 SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
 Residue Identity = 54% Matches = 6 Mismatches = 5
 Gaps = 0 Conservative Substitutions = 0

X 10
 GGQPLWITATK
 || | |||

GPSRLAWSWFVIALVASLGGANVATAGLLDLNDVPAWLRILVAAMPALAFMGGTLLAHTATHHEPEAPAPTQ
 70 80 90 100 110 X 120 X 130

APEPPAFTDEHDLVRVDDTEEPPELPAPGLQAPAPPAV
 140 150 160 170

11. US-08-249-182-8 (1-11)

A44914 phosphate-dependent exoribonuclease - Bacillus sub

ENTRY A44914 #type complete
 TITLE phosphate-dependent exoribonuclease - Bacillus subtilis
 ORGANISM #formal_name Bacillus subtilis
 DATE 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 17-Feb-1994
 ACCESSIONS A44914
 REFERENCE A44914
 #authors Craven, M.G.; Henner, D.J.; Alessi, D.; Schauer, A.T.; Ost, K.A.; Deutscher, M.P.; Friedman, D.I.
 #journal J. Bacteriol. (1992) 174:4727-4735
 #title Identification of the rph (RNase PH) gene of Bacillus subtilis: evidence for suppression of cold-sensitive mutations in Escherichia coli.
 #cross-references MUID:92325065
 #accession A44914
 ##status preliminary
 ##molecule_type DNA
 ##residues 1-245 ##label CRA
 ##cross-references NCBI:P108178
 ##note sequence extracted from NCBI backbone

GENETICS
 #gene rph

SUMMARY #length 248 #molecular-weight 26539 #checksum 7992
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                GG@PLWITATK
                                ||  |||
HDGRGHDELRPITFDLDFISHPEGSVLITAGNTKVICNASVEDRVPFRLGGGKGWITAEYSMLPRATNGRT
   10      20      30      40      50 X      60 X      70

IRESSKKGKISGRTEI@RLIGRALRAVVDLEKLGERTIW
   80      90     100     110
```

12. US-08-249-182-8 (1-11)

S03888 photosystem II oxygen-evolving complex protein 2 p

ENTRY S03888 #type fragment
TITLE photosystem II oxygen-evolving complex protein 2 precursor -
white mustard (fragment)
ALTERNATE_NAMES photosystem II extrinsic membrane protein 23K chain
ORGANISM #formal_name Sinapis alba #common_name white mustard
DATE 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
31-Dec-1993
ACCESSIONS S03888
REFERENCE S03888
#authors Wenng, A.; Ehmann, B.; Schaefer, E.
#journal FEBS Lett. (1989) 246:140-144
#title The 23 kDa polypeptide of the photosynthetic oxygen-evolving
complex from mustard seedlings (Sinapis alba L.).
Nucleotide sequence of cDNA and evidence for phytochrome
control of its mRNA abundance.
#cross-references MUID:89211386
#accession S03888
##molecule_type mRNA
##residues 1-248 ##label WEN
##cross-references EMBL:Y07498
CLASSIFICATION #superfamily photosystem II oxygen-evolving complex protein 2
KEYWORDS chloroplast; membrane protein; photosynthesis; photosystem
II; thylakoid
FEATURE
1-62 #domain transit peptide (fragment) #label TNP\
63-248 #protein photosystem II oxygen-evolving complex protein
2 #label MAT
SUMMARY #length 248 #checksum 446
SEQUENCE

Initial Score = 6 Optimized Score = 6 Significance = 4.08
Residue Identity = 54% Matches = 6 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

```

                                X      10
                                GG@PLWITATK
                                ||  |||
QAYFGETASEGGFDNNAVATANILETNI@DVGGKPPYYLSVLTRTADGDEGGKHQLITATVNGGKLYICKA@
   160     170     180     190     200 X     210 X     220

AGDKRWFKGANKFVEKAATSFSVA
   230     240
```

13. US-08-249-182-8 (1-11)

S22763 photosystem II oxygen-evolving complex protein 2 p

```

##status      preliminary
##molecule_type  protein
##residues    1-114 ##label STR
##cross-references  NCBIP:78526; NCBIP:78523; NCBIP:78521; NCBIP:78518;
                   NCBIP:78515; NCBIP:78512; NCBIP:78511; NCBIP:78510;
                   NCBIP:78509; NCBIP:78508; NCBIP:78503
##note        sequence extracted from NCBI backbone
SUMMARY       #length 114 #checksum 7335
SEQUENCE

```

```

Initial Score   =    16  Optimized Score =    16  Significance = 13.09
Residue Identity = 100%  Matches         =    16  Mismatches  =    0
Gaps            =    0  Conservative Substitutions =    0

```

```

                X      10      X
                VNSMGTVFVVGYPGTFK
                |||||
TEFLSNYLTNVDDITLVPGLGRDIEHLTSLDFFRVNSMGTVFVVGYPGTFKGGQPLWITATKSPPFENINLY
      10      20      30      X  40      50      60      70

YDVPWNETIPEEVTXPNYLQAEVSYPAFK
      80      90     100

```

2. US-08-249-182-9 (1-16)

A39216 plasma cell membrane protein PC-1 - human

```

ENTRY      A39216      #type complete
TITLE      plasma cell membrane protein PC-1 - human
ORGANISM    #formal_name Homo sapiens #common_name man
DATE        23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change
              31-Dec-1993
ACCESSIONS  A39216
REFERENCE   A39216
#authors    Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.;
              Goding, J.W.
#journal     J. Biol. Chem. (1990) 265:17506-17511
#title       Plasma cell membrane glycoprotein PC-1. cDNA cloning of the
              human molecule, amino acid sequence, and chromosomal
              location.
#cross-references  MUID:91009202
#accession  A39216
##status    preliminary
##molecule_type  mRNA
##residues  1-925 ##label BUC
##cross-references  GB:J05654
KEYWORDS    membrane protein
SUMMARY     #length 925 #molecular-weight 104924 #checksum 7446
SEQUENCE

```

```

Initial Score   =    10  Optimized Score =    11  Significance = 7.48
Residue Identity = 70%  Matches         =    12  Mismatches  =    4
Gaps            =    1  Conservative Substitutions =    0

```

```

                X      10      X
                VNS-MGTVFVVGYPGTFK
                ||| ||| ||| |||
YLKHF LPKRLHFAKSDRIEPLTFYLD PQWQLALNP SERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEA
490      500      510      520      530      540      550      X 560

DTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPK
      570      580      590      600

```

3. US-08-249-182-9 (1-16)

S21706 nucleotide pyrophosphatase - human

ENTRY S21706 #type complete
 TITLE nucleotide pyrophosphatase - human
 ALTERNATE_NAMES plasma cell membrane protein PC-1
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
 18-Jun-1993
 ACCESSIONS S21706; S23587
 REFERENCE S21706
 #authors Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.;
 Kobayashi, H.; Inoue, T.; Suzuki, H.; Fukui, S.; Tsukahara,
 M.; Kajii, T.; Yamashina, I.
 #journal Arch. Biochem. Biophys. (1992) 295:180-187
 #title Molecular cloning of cDNAs for human fibroblast nucleotide
 pyrophosphatase.
 #cross-references MUID:92246539
 #accession S21706
 ##molecule_type mRNA
 ##residues 1-925 ##label FUN1
 ##note sequence not compared to nucleotide translation
 #accession S23587
 ##molecule_type protein
 ##residues 116-121;247-271,'X',273-275;279-280,'X',282-283;303-316;
 362-364;449-465;482-525;529-534,'X',536-551,'X',553,
 'X',555-556;597-606;'X',727-730;775-782;840-846,'XX',
 849-852,'X',854-859 ##label FUN2
 ##note it is uncertain whether Met-1 or Met-53 is the initiator

GENETICS

#map_position 6q22-q23

KEYWORDS glycoprotein; membrane protein

FEATURE

77-97 #domain transmembrane #status predicted #label TMM\
 179,285,341,477,
 578,585,643,700,
 731,748 #binding_site carbohydrate (Asn) (covalent) #status
 predicted

SUMMARY #length 925 #molecular-weight 104924 #checksum 7446

SEQUENCE

Initial Score = 10 Optimized Score = 11 Significance = 7.48
 Residue Identity = 70% Matches = 12 Mismatches = 4
 Gaps = 1 Conservative Substitutions = 0

```

                                X      10      X
                                VNS-MQTVFVGYGPTFK
                                | | | | | | | |
YLKHF LPKRLHFAKSDRIEPLTFYLD P Q W Q L A L N P S E R K Y C G S G F H G S D N V F S N M Q A L F V G Y G P G F K H G I E A
490      500      510      520      530      540      550      X 560

DTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPK
      570      580      590      600

```

4. US-08-249-182-9 (1-16)

A27410 plasma cell membrane protein PC-1 - mouse

ENTRY A27410 #type complete
 TITLE plasma cell membrane protein PC-1 - mouse
 ORGANISM #formal_name Mus musculus #common_name house mouse
 DATE 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
 31-Dec-1993
 ACCESSIONS A27410
 REFERENCE A27410
 #authors van Driel, I.R.; Goding, J.W.
 #journal J. Biol. Chem. (1987) 262:4882-4887
 #title Plasma cell membrane glycoprotein PC-1. Primary structure

deduced from cDNA clones.

#cross-references MUID:87165906

#accession A27410

##molecule_type mRNA

##residues 1-905 ##label VAN

##note the authors translated the codon CAG for residue 24 as
Glu

KEYWORDS membrane protein

SUMMARY #length 905 #molecular-weight 102880 #checksum 1749

SEQUENCE

Initial Score = 9 Optimized Score = 9 Significance = 6.55
Residue Identity = 56% Matches = 9 Mismatches = 7
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                || | ||| ||
LKPFLPKRLHFAKSDRIEPLTFYLDPQWLALNPSEKCYGSGFHGSDNLFNMQALFIGYGPAPFKHGAEDV
 480      490      500      510      520 X      530      540

SFENIEVYNLMCDLLGLIPAPNNGSHGSLNHLLKKPIYNPSHPK
 550      560      570      580
```

5. US-08-249-182-9 (1-16)

RKAALC ribulose-bisphosphate carboxylase (EC 4.1.1.39) la

ENTRY RKAALC #type complete

TITLE ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain
precursor - alfalfa chloroplast

ORGANISM #formal_name chloroplast Medicago sativa #common_name alfalfa

DATE 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change
30-Jun-1993

ACCESSIONS A25578

REFERENCE A25578

#authors Aldrich, J.; Cherney, B.; Merlin, E.; Palmer, J.

#journal Nucleic Acids Res. (1986) 14:9535

#title Sequence of the rbcL gene for the large subunit of ribulose
bisphosphate carboxylase-oxygenase from alfalfa.

#cross-references MUID:87091586

#accession A25578

##molecule_type protein

##residues 1-474 ##label ALD

##cross-references GB:X04975

COMMENT In addition to Lys-201, another lysine, it is not certain which,
may be the site of autocatalytic carbamylation.

GENETICS

#gene rbcL

#genome chloroplast

CLASSIFICATION #superfamily ribulose-bisphosphate carboxylase large chain

KEYWORDS Calvin cycle; carbon dioxide fixation; carbon-carbon lyase;
carboxy-lyase; chloroplast; monooxygenase;
photorespiration; photosynthesis

FEATURE

3-474 #protein ribulose-bisphosphate carboxylase large chain

#status predicted #label MAT\

175,334 #active_site Lys (ribulose-bisphosphate-binding) #status
predicted\

201 #binding_site carbon dioxide (Lys) (covalent) (by
Rubisco activase) #status predicted\

203 #binding_site magnesium (Asp) #status predicted

SUMMARY #length 474 #molecular-weight 52625 #checksum 2374

SEQUENCE

Initial Score = 8 Optimized Score = 8 Significance = 5.61

Residue Identity = 50% Matches = 8 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                || | | || ||
TGTWTTVWTDGLTSLDRYKGRCYHIEPVAGEETQFIAYVAYPLDLFEEGSVNYMFTSIVGNVFGFKALRALR
   70      80      90      100      110 X      120      130

LEDLRIPAAYVKTFQGPPEGIQVERDKLNKYGRPLLGCITKPKL
   140      150      160      170
```

6. US-08-249-182-9 (1-16)

DSNVAC superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Autog

ENTRY DSNVAC #type complete
TITLE superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Autographa
californica nuclear polyhedrosis virus (strain L1)
ORGANISM #formal_name Autographa californica nuclear polyhedrosis
virus, AcMNPV
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
24-Feb-1994
ACCESSIONS A40564
REFERENCE A40564
#authors Tomalski, M.D.; Eldridge, R.; Miller, L.K.
#journal Virology (1991) 184:149-161
#title A baculovirus homolog of a Cu/Zn superoxide dismutase gene.
#cross-references MUID:91335744
#accession A40564
##molecule_type DNA
##residues 1-151 ##label TOM
##cross-references GB:M68862
CLASSIFICATION #superfamily superoxide dismutase (Cu-Zn)
KEYWORDS oxidoreductase
SUMMARY #length 151 #molecular-weight 16182 #checksum 4639
SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                || | | |||
HEYGDTSNGCTSAGEHFNPNTNEDHGAPDAEIRHVGD LGNIKSAGYNSLTEVNMDNVMSLYGPHNIIGRSLV
   50      60      70      80      90 X      100      110

VHTDKDDLGLTDHPLSKTTGNSGGRLGCGIIAICK
   120      130      140      150
```

7. US-08-249-182-9 (1-16)

A35216 FPD4 protein - fowlpox virus (strain FP-1)

ENTRY A35216 #type complete
TITLE FPD4 protein - fowlpox virus (strain FP-1)
ORGANISM #formal_name fowlpox virus
DATE 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change
30-Sep-1993
ACCESSIONS A35216
REFERENCE A35216
#authors Tartaglia, J.; Winslow, J.; Goebel, S.; Johnson, G.P.;
Taylor, J.; Paoletti, E.
#journal J. Gen. Virol. (1990) 71:1517-1524
#title Nucleotide sequence analysis of a 10.5 kbp HindIII fragment

of fowlpox virus: relatedness to the central portion of the
vaccinia virus HindIII D region.

#cross-references MUID:90324937

#accession A35216

##status preliminary

##molecule_type DNA

##residues 1-218 ##label TAR

##cross-references GB:X17202

CLASSIFICATION #superfamily vaccinia virus D4 protein

SUMMARY #length 218 #molecular-weight 25563 #checksum 5203

SEQUENCE

Initial Score = 7 Optimized Score = 8 Significance = 4.68
Residue Identity = 50% Matches = 8 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQTVEFVGYGPTFK
                                || || || |
NYYLSCREGEAKSHKIFWERLADVFINHIAAYVSFVFLGKSDFSNFRSILNSPTTVVVGYPAAARNRQFDT
120      130      140      150      160      170      180      X  190

DETFEIVNTLLELKNEPRINWVGFEI
      200      210
```

8. US-08-249-182-9 (1-16)

A24637 T-cell surface glycoprotein CD8 precursor - rat

ENTRY A24637 #type complete

TITLE T-cell surface glycoprotein CD8 precursor - rat

ALTERNATE_NAMES MRC OX-8 antigen

ORGANISM #formal_name Rattus norvegicus #common_name Norway rat

DATE 17-Sep-1987 #sequence_revision 17-Sep-1987 #text_change
31-Dec-1993

ACCESSIONS A24637

REFERENCE A24637

#authors Johnson, P.; Gagnon, J.; Barclay, A.N.; Williams, A.F.

#journal EMBO J. (1985) 4:2539-2545

#title Purification, chain separation and sequence of the MRC OX-8
antigen, a marker of rat cytotoxic T lymphocytes.

#cross-references MUID:86030231

#accession A24637

##molecule_type mRNA

##residues 1-236 ##label JOH

CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology

KEYWORDS glycoprotein

SUMMARY #length 236 #molecular-weight 26196 #checksum 3912

SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQTVEFVGYGPTFK
                                ||| | | ||
DPNLFSARKENNKYILTSKFSTKNQGYFCSITSNSVMYFSPLVPVFQKVNIIITKPVTRAPTVPVPPPTGT
90      100      110      120      130      140      150      X  160

PRPLRPEACRPGASGSVEGMGLGFACDIYIWAPLAGICAVLLLS
      170      180      190      200
```

9. US-08-249-182-9 (1-16)

A45442 Sec13p=endoplasmic reticulum vesicle formation pro

ENTRY A45442 #type complete
 TITLE Sec13p=endoplasmic reticulum vesicle formation protein - yeast (Saccharomyces cerevisiae)
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 21-Sep-1993; #sequence_revision 21-Sep-1993; #text_change 21-Sep-1993
 ACCESSIONS A45442
 REFERENCE A45442
 #authors Pryer, N.K.; Salama, N.R.; Schekman, R.; Kaiser, C.A.
 #journal J. Cell Biol. (1993) 120:865-875
 #title Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro.
 #cross-references MUID:93163112
 #accession A45442
 ##status preliminary
 ##molecule_type nucleic acid
 ##residues 1-297 ##label PRY
 ##cross-references NCBIN:124845; NCBIP:124846
 ##note sequence extracted from NCBI backbone
 SUMMARY #length 297 #molecular-weight 33043 #checksum 717
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                ||| |      |||
HEGPVWRVDWAHPKFGTILASCSYDGKVLWKEENGRWSQIAVHAVHSASVNSVQWAPHEYGPLLLVASSDG
      60          70          80          90         100 X      110         120

KVSVEFEKNGTTSPIIIDAHAIGVNSASWAPATIEEDGEHNGT
      130         140         150         160

```

10. US-08-249-182-9 (1-16)

S30803 SEC13 protein - yeast (Saccharomyces cerevisiae)

ENTRY S30803 #type complete
 TITLE SEC13 protein - yeast (Saccharomyces cerevisiae)
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 28-May-1993 #sequence_revision 28-May-1993 #text_change 06-May-1994
 ACCESSIONS S30803
 REFERENCE S30803
 #authors Pryer, N.K.; Salama, N.R.; Schekman, R.; Kaiser, C.A.
 #submission submitted to the EMBL Data Library, November 1992
 #description Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro.
 #accession S30803
 ##molecule_type DNA
 ##residues 1-297 ##label PRY
 ##cross-references EMBL:L05929
 GENETICS
 #gene LISTA:SEC13
 SUMMARY #length 297 #molecular-weight 33043 #checksum 717
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK

```


HEGPVWRVDWAHPKFGTILASCSYDGKVLWKEENGRWSQIAVHAVHSASVNSVQWAPHEYGPLLLVASSDG
60 70 80 90 100 X 110 120

KVSVEFKENGTTSPIIIDAHAGVNSASWAPATIEEDGEHNGT
130 140 150 160

11. US-08-249-182-9 (1-16)

S06938 glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - uni

ENTRY S06938 #type complete
TITLE glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - unidentified
bacterium
ORGANISM #formal_name unidentified bacterium
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
31-Dec-1993
ACCESSIONS S06938
REFERENCE S06938
#authors Cock, J.M.; Schmidt, R.R.
#journal Nucleic Acids Res. (1989) 17:10500
#title A glutamate dehydrogenase gene sequence.
#cross-references MUID:90098893
#accession S06938
##molecule_type DNA
##residues 1-446 ##label CDC
##cross-references EMBL:X16399
##note the translation of the nucleotide sequence is not given
in this paper
CLASSIFICATION #superfamily glutamate dehydrogenase (NAD(P)+)
KEYWORDS NADP; oxidoreductase
SUMMARY #length 446 #molecular-weight 48490 #checksum 8711
SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

X 10 X
VNSMQTVFVGYGPTFK
|| || ||
LERLVEPERIIQFRVSWVDDRQGVQVNRAFRVQFNSAIGPYKGGMRHPSVNLKILKFLGFETFKNALTTL
60 70 80 90 100 110 X 120
PMGGGKGGSDFDPKGKSQGRIMRFCQALMTLYRHLGPDTDVPA
130 140 150 160

12. US-08-249-182-9 (1-16)

A33504 glutamate dehydrogenase (EC 1.4.1.2) - Salmonella

ENTRY A33504 #type complete
TITLE glutamate dehydrogenase (EC 1.4.1.2) - Salmonella typhimurium
ORGANISM #formal_name Salmonella typhimurium
DATE 08-Dec-1989 #sequence_revision 08-Dec-1989 #text_change
30-Sep-1993
ACCESSIONS A33504
REFERENCE A33504
#authors Bansal, A.; Dayton, M.A.; Zalkin, H.; Colman, R.F.
#journal J. Biol. Chem. (1989) 264:9827-9835
#title Affinity labeling of a glutamyl peptide in the coenzyme
binding site of NADP(+)-specific glutamate dehydrogenase of
Salmonella typhimurium by 2-[4-bromo-2,
3-dioxobutyl]thiol-1,N(6)-ethenoadenosine 2',
5'-bisphosphate.
#cross-references MUID:8925551

#accession A33304
 ##status preliminary
 ##molecule_type DNA
 ##residues 1-447 ##label BAN
 ##cross-references GB:M24021; GB:J04814
 CLASSIFICATION #superfamily glutamate dehydrogenase (NAD(P)+)
 KEYWORDS oxidoreductase
 SUMMARY #length 447 #molecular-weight 48574 #checksum 5173
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||      ||      |||
LERLVEPERVIQFRVVWLDKNGVQVNRARVQFNSAIGPYKGGMRFHPSVNL SILKFLGFETFKNALTTL
      60      70      80      90      100      110      X 120

PMGGGKGGSDFDPKGKSEGEVMRFCQALMTELYRHLGPD TDVPA
      130      140      150      160
  
```

13. US-08-249-182-9 (1-16)

A22413 glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - Esc

ENTRY A22413 #type complete
 TITLE glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - Escherichia coli
 ALTERNATE_NAMES NADP-specific glutamate dehydrogenase
 ORGANISM #formal_name Escherichia coli
 DATE 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 31-Dec-1993
 ACCESSIONS A22413
 REFERENCE A22413
 #authors Valle, F.; Becerril, B.; Chen, E.; Seeburg, P.; Heyneker, H.; Bolivar, F.
 #journal Gene (1984) 27:193-199
 #title Complete nucleotide sequence of the glutamate dehydrogenase gene from Escherichia coli K-12.
 #cross-references MUID:84209849
 #contents Strain K12
 #accession A22413
 ##molecule_type DNA
 ##residues 1-447 ##label VAL
 CLASSIFICATION #superfamily glutamate dehydrogenase (NAD(P)+)
 KEYWORDS NADP; oxidoreductase
 SUMMARY #length 447 #molecular-weight 48581 #checksum 2843
 SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQT V F V G Y G P T F K
                                ||      ||      |||
LERLVEPERVIQFRVVWVDDRNGIQVNRARVQFSSAIGPYKGGMRFHPSVNL SILKFLGFETFKNALTTL
      60      70      80      90      100      110      X 120

PMGGGKGGSDFDPKGKSEGEVMRFCQALMTELYRHLGADTDVPA
      130      140      150      160
  
```

14. US-08-249-182-9 (1-16)

DEECEN glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - ESE

ENTRY DEECEN #type complete
TITLE glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - Escherichia coli
ALTERNATE_NAMES NADP-specific glutamate dehydrogenase
ORGANISM #formal_name Escherichia coli
DATE 03-Aug-1984 #sequence_revision 20-Sep-1984 #text_change 31-Dec-1993
ACCESSIONS A00382
REFERENCE A00382
#authors McPherson, M.J.; Wootton, J.C.
#journal Nucleic Acids Res. (1983) 11:5257-5266
#title Complete nucleotide sequence of the Escherichia coli gdhA gene.
#cross-references MUID:83272967
#accession A00382
##molecule_type DNA
##residues 1-447 ##label MCP
COMMENT This enzyme is a hexamer of identical chains.
GENETICS
#gene gdhA
#map_position 27 min
CLASSIFICATION #superfamily glutamate dehydrogenase (NAD(P)+)
KEYWORDS oxidoreductase
SUMMARY #length 447 #molecular-weight 48581 #checksum 2843
SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMØTVFVGYGPTFK
                                ||      ||      |||
LERLVEPERVIQFRVVWVDDRNGIQVNRAWRVQFSSAIGPYKGGMRFHPSVNLKFLGFETFKNALTTL
      60          70          80          90         100         110      X 120

PMGGGKGGSDFDPKGKSEGEVMRFGALMTELYRHLGADTDVPA
      130          140          150          160
```

15. US-08-249-182-9 (1-16)

APBOL leucyl aminopeptidase (EC 3.4.11.1) - bovine

ENTRY APBOL #type complete
TITLE leucyl aminopeptidase (EC 3.4.11.1) - bovine
ALTERNATE_NAMES cytosol aminopeptidase
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 31-Aug-1980 #sequence_revision 31-Aug-1980 #text_change 31-Dec-1993
ACCESSIONS A00907
REFERENCE A92380
#authors Cuypers, H.T.; van Loon-Klaassen, L.A.H.; Vree Egberts, W.T.M.; de Jong, W.W.; Bloemendal, H.
#journal J. Biol. Chem. (1982) 257:7077-7085
#title The primary structure of leucine aminopeptidase from bovine eye lens.
#cross-references MUID:82213853
#accession A00907
##molecule_type protein
##residues 1-478 ##label CUY
REFERENCE A92381
#authors Cuypers, H.T.; van Loon-Klaassen, L.A.H.; Vree Egberts, W.T.M.; de Jong, W.W.; Bloemendal, H.
#journal J. Biol. Chem. (1982) 257:7086-7091

#title sulfhydryl content of bovine eye lens leucine aminopeptidase.
 Determination of the reactivity of the sulfhydryl groups of
 the zinc metalloenzyme, of the enzyme activated by Mg(2+),
 Mn(2+), and Co(2+), and of the metal-free apoenzyme.

#cross-references MUID:82213854

#contents annotation

#note no disulfide bonds are present

COMMENT This protein, isolated from calf lens, is a hexamer of identical
 chains.

CLASSIFICATION #superfamily cytosol aminopeptidase

KEYWORDS alpha-aminoacylpeptide hydrolase

SUMMARY #length 478 #molecular-weight 51691 #checksum 5529

SEQUENCE

Initial Score = 7 Optimized Score = 7 Significance = 4.68
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMQTVFVGYGPTFK
                                |||| | ||
NLKSASIKTDVFI RPKSWIEE QENG SFLSVAKGSEPPVFLEIHYKGSPNASEPPLVFVGKGITFD SGGISI
190      200      210      220      230      240      250      X 260
```

```

KAAANMDLMRADMGGAATICS AIVSAAKLDLPINIVGLAPLCEN
      270      280      290      300
```

> 0 <
0| |0 IntelliGenetics
> 0 <

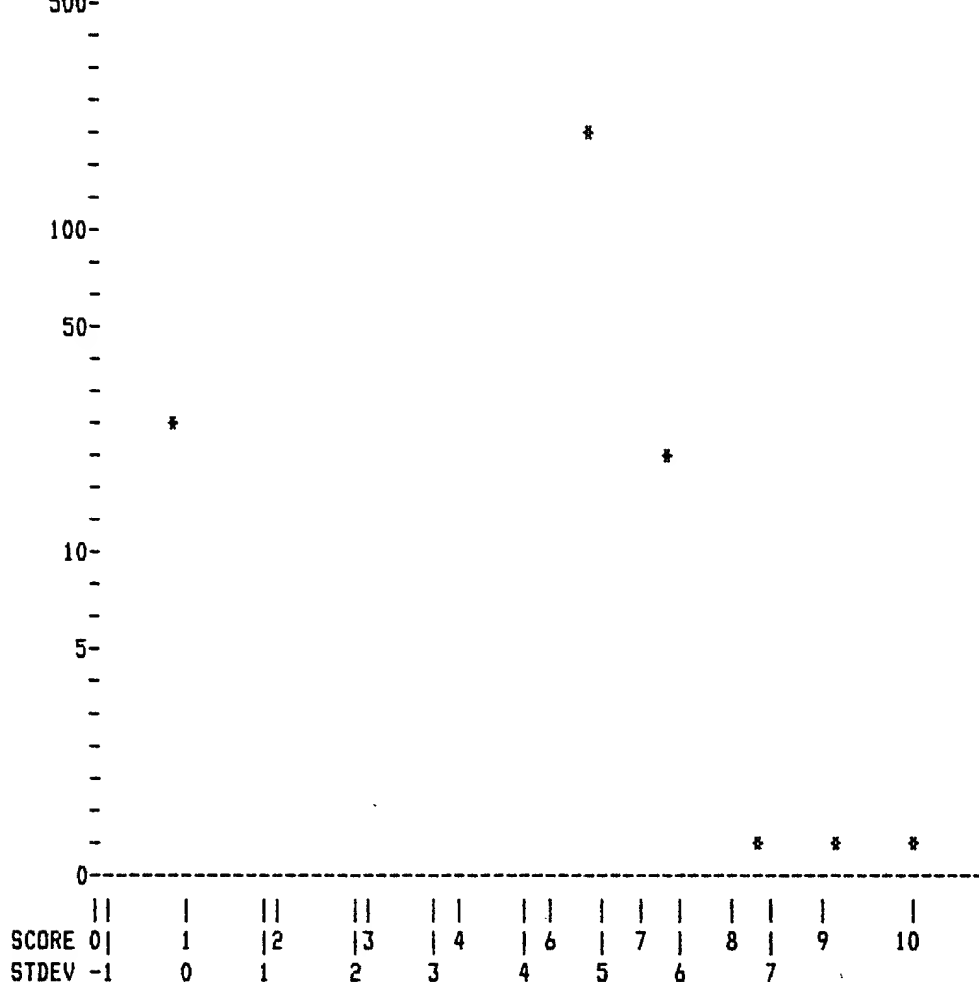
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file u249_9s.res made by on Thu 22 Sep 94 10:40:14-PDT.

Query sequence being compared:US-08-249-182-9 (1-16)
Number of sequences searched: 36000
Number of scores above cutoff: 4370

Results of the initial comparison of US-08-249-182-9 (1-16) with:
Data bank : Swiss-Prot 28, all entries

100000-
-
N -
U50000-
M -
B -
E -
R - *
-
O -
F10000- *
-
S -
E 5000- *
Q -
U -
E -
N -
C *
E - *
S 1000-
-
-



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	5
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		
Initial scores to save	40	Alignments to save	15
Optimized scores to save	0	Display context	50

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	2	3	0.96
Times:	CPU	Total Elapsed	
	00:00:51.91	00:00:58.00	

Number of residues:	12496420
Number of sequences searched:	36000
Number of scores above cutoff:	4370

Cut-off raised to 2.
 Cut-off raised to 3.
 Cut-off raised to 4.
 Cut-off raised to 5.

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 8 standard deviations above mean ****						
1. PC1_HUMAN	PLASMA-CELL MEMBRANE GLYCOPRO	873	10	11	8.29	0
**** 7 standard deviations above mean ****						
2. PC1_MOUSE	PLASMA-CELL MEMBRANE GLYCOPRO	871	9	9	7.26	0
**** 6 standard deviations above mean ****						
3. RBL_MEDSA	RIBULOSE BISPHOSPHATE CARBOXY	474	8	8	6.22	0
**** 5 standard deviations above mean ****						
4. SODC_NPVAC	PUTATIVE SUPEROXIDE DISMUTASE	151	7	7	5.18	0
5. UNG_FOWP1	URACIL-DNA GLYCOSYLASE (EC 3.	218	7	8	5.18	0
6. CD8A_RAT	T-CELL SURFACE GLYCOPROTEIN C	236	7	7	5.18	0
7. SC13_YEAST	PROTEIN TRANSPORT PROTEIN SEC	297	7	7	5.18	0
8. DHE4_UNKP	NADP-SPECIFIC GLUTAMATE DEHYD	446	7	7	5.18	0
9. DHE4_SALTY	NADP-SPECIFIC GLUTAMATE DEHYD	447	7	7	5.18	0
10. DHE4_ECOLI	NADP-SPECIFIC GLUTAMATE DEHYD	447	7	7	5.18	0
11. AMPL_BOVIN	CYTOSOL AMINOPEPTIDASE (EC 3.	478	7	7	5.18	0
12. HEMA_PI3B	HEMAGGLUTININ-NEURAMINIDASE (572	7	7	5.18	0
13. NAK1_HUMAN	EARLY RESPONSE PROTEIN NAK1.	598	7	7	5.18	0
14. TOXA_PSEAE	EXOTOXIN A PRECURSOR (NAD-DEP	638	7	7	5.18	0
15. FEPA_ECOLI	FERRIC ENTEROCHELIN RECEPTOR	745	7	7	5.18	0
16. VG43_HSVI1	HYPOTHETICAL GENE 43 PROTEIN.	891	7	7	5.18	0
17. VGL2_IBVM	E2 GLYCOPROTEIN PRECURSOR (SP	1162	7	7	5.18	0
18. VGL2_IBVK	E2 GLYCOPROTEIN PRECURSOR (SP	1162	7	7	5.18	0
19. TYK2_HUMAN	NON-RECEPTOR TYROSINE-PROTEIN	1187	7	7	5.18	0
20. VOR1_FXMV	152 KD PROTEIN (DRF 1).	1335	7	7	5.18	0
21. MPRI_BOVIN	CATION-INDEPENDENT MANNOSE-6-	2499	7	7	5.18	0
**** 4 standard deviations above mean ****						
22. Y601_BPT4	HYPOTHETICAL 8.5 KD PROTEIN I	70	6	6	4.15	0
23. YVDB_VACCV	HYPOTHETICAL 8.5 KD PROTEIN.	80	6	6	4.15	0
24. YVDB_VACCC	HYPOTHETICAL 8.5 KD PROTEIN.	80	6	6	4.15	0
25. Y242_BPT4	HYPOTHETICAL 11.0 KD PROTEIN	92	6	6	4.15	0
26. PLAS_PETCR	PLASTOCYANINS A AND B.	97	6	7	4.15	0
27. PLAS_ORYSA	PLASTOCYANIN.	97	6	6	4.15	0
28. PLAS_CUCSA	PLASTOCYANIN.	99	6	7	4.15	0
29. ISS_ECOLI	HYPOTHETICAL ISS PROTEIN.	102	6	6	4.15	0
30. K2M1_SHEEP	KERATIN, TYPE II MICROFIBRILL	109	6	6	4.15	0
31. HV35_MOUSE	IG HEAVY CHAIN V-III REGION (111	6	6	4.15	0
32. HV34_MOUSE	IG HEAVY CHAIN V REGION (AMPC	113	6	6	4.15	0
33. HV31_MOUSE	IG HEAVY CHAIN V-III REGION (113	6	6	4.15	0
34. HV30_MOUSE	IG HEAVY CHAIN V-III REGION (113	6	6	4.15	0
35. HV29_MOUSE	IG HEAVY CHAIN V-III REGION (113	6	6	4.15	0
36. HV28_MOUSE	IG HEAVY CHAIN V-III REGION (113	6	6	4.15	0
37. HV27_MOUSE	IG HEAVY CHAIN V-III REGION (113	6	6	4.15	0
38. HV33_MOUSE	IG HEAVY CHAIN V-III REGION (115	6	6	4.15	0
39. HV32_MOUSE	IG HEAVY CHAIN V-III REGION (115	6	6	4.15	0
40. SFP1_BOVIN	SEMINAL PLASMA PROTEIN PDC-10	134	6	6	4.15	0

1. US-08-249-182-9 (1-16)

PC1_HUMAN PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE P

ID PC1_HUMAN STANDARD; PRT; 873 AA.
AC P22413;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE PHOSPHODIESTERASE I
DE (EC 3.1.4.1) / NUCLEOTIDE PYROPHOSPHATASE (EC 3.6.1.9)).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

UC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91009202
 RA BUCKLEY M.F., LOVELAND K.A., MCKINSTRY W.J., GARSON O.M., GODING J.W.;
 RL J. BIOL. CHEM. 265:17506-17511(1990).
 CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.
 CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES
 CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED
 CC OLIGO-NUCLEOTIDES.
 CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER
 CC OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE
 CC OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.
 CC -!- SIMILARITY: CONTAINS TWO TANDEM COPIES OF A SOMATOMEDIN-B TYPE
 CC DOMAIN.
 DR EMBL; M57736; HSPC1Q1.
 DR PIR; A39216; A39216.
 DR MIM; 173335; TENTH EDITION.
 DR PROSITE; PS00524; SOMATOMEDIN_B.
 KW GLYCOPROTEIN; TRANSMEMBRANE; DUPLICATION; SIGNAL-ANCHOR; HYDROLASE.
 FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 25 45 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 46 873 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 52 92 SOMATOMEDIN-B LIKE.
 FT DOMAIN 93 136 SOMATOMEDIN-B LIKE.
 FT CARBOHYD 127 127 POTENTIAL.
 FT CARBOHYD 233 233 POTENTIAL.
 FT CARBOHYD 289 289 POTENTIAL.
 FT CARBOHYD 425 425 POTENTIAL.
 FT CARBOHYD 533 533 POTENTIAL.
 FT CARBOHYD 591 591 POTENTIAL.
 FT CARBOHYD 648 648 POTENTIAL.
 FT CARBOHYD 679 679 POTENTIAL.
 FT CARBOHYD 696 696 POTENTIAL.
 SQ SEQUENCE 873 AA; 99929 MW; 4095996 CN;

Initial Score = 10 Optimized Score = 11 Significance = 8.29
 Residue Identity = 70% Matches = 12 Mismatches = 4
 Gaps = 1 Conservative Substitutions = 0

X 10 X
 VNS-MQTVFVGYGPTFK
 | | | | | | | |
 YLKHFLPKRLHFAKSDRIEPLTFYLDPOWLALNPSEKRYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEA
 440 450 460 470 480 490 500 X

DTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPK
 510 520 530 540 550

2. US-08-249-182-9 (1-16)

PC1_MOUSE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE P

ID PC1_MOUSE STANDARD; PRT; 871 AA.
 AC P06802;
 DT 01-JAN-1988 (REL. 06, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 (ALKALINE PHOSPHODIESTERASE I
 DE (EC 3.1.4.1) / NUCLEOTIDE PYROPHOSPHATASE (EC 3.6.1.9)) (LY-41).
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.

RN [1]
 RP SEQUENCE FROM N.A.
 RM 87165906
 RA VAN DRIEL I.R., GODING J.W.;
 RL J. BIOL. CHEM. 262:4882-4887(1987).
 RN [2]
 RP PARTIAL SEQUENCE.
 RM 85056299
 RA STEARNE P.A., VAN DRIEL I.R., GREGO B., SIMPSON R.J., GODING J.W.;
 RL J. IMMUNOL. 134:443-448(1985).
 RN [3]
 RP FUNCTION, AND SEQUENCE FROM N.A.
 RM 91271356
 RA REBBE N.F., TONG B.D., FINLEY E.M., HICKMAN S.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:5192-5196(1991).
 CC -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.
 CC -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES
 CC SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED
 CC OLIGO-NUCLEOTIDES.
 CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON THE SURFACE OF
 CC ANTIBODY-SECRETING CELLS.
 CC -!- SIMILARITY: CONTAINS TWO TANDEM COPIES OF A SOMATOMEDIN-B TYPE
 CC DOMAIN.
 DR EMBL; J02700; MMPC1B.
 DR PIR; A27410; A27410.
 DR PROSITE; PS00524; SOMATOMEDIN_B.
 KW GLYCOPROTEIN; TRANSMEMBRANE; DUPLICATION; SIGNAL-ANCHOR; HYDROLASE.
 FT NOD_RES 21 21 BLOCKED.
 FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 25 45 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 46 871 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 52 92 SOMATOMEDIN-B LIKE.
 FT DOMAIN 93 136 SOMATOMEDIN-B LIKE.
 FT CARBOHYD 127 127 POTENTIAL.
 FT CARBOHYD 233 233 POTENTIAL.
 FT CARBOHYD 289 289 POTENTIAL.
 FT CARBOHYD 425 425 POTENTIAL.
 FT CARBOHYD 533 533 POTENTIAL.
 FT CARBOHYD 590 590 POTENTIAL.
 SQ SEQUENCE 871 AA; 99487 MW; 4094440 CN;

Initial Score = 9 Optimized Score = 9 Significance = 7.26
 Residue Identity = 56% Matches = 9 Mismatches = 7
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSMGTVFVGYGPTFK
 II I IIII II
 LKPF LPKRLHFAKSDRIEPLTFYLD P Q W L A L N P S E R K Y C G S G F H G S D N L F S N M Q A L F I G Y G P A F K H G A E V D
 440 450 460 470 480 490 500 X 510
 S F E N I E V Y N L M C D L L G L I P A P N N G S H G S L N H L L K K P I Y N P S H P K
 520 530 540 550

3. US-08-249-182-9 (1-16)

RBL_MEDSA RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PREC

ID RBL_MEDSA STANDARD; PRT; 474 AA.
 AC P04991;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)

DE RIBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.397).
 GN RBCL.
 OS MEDICAGO SATIVA (ALFALFA).
 OG CHLOROPLAST.
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
 OC FABACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. REGEN S;
 RM 87091586
 RA ALDRICH J., CHERNEY B., MERLIN E., PALMER J.;
 RL NUCLEIC ACIDS RES. 14:9535-9535(1986).
 CC -!- FUNCTION: RUBISCO, A MAJOR COMPONENT OF LEAF PROTEIN, CATALYSES
 CC TWO REACTIONS: THE CARBOXYLATION OF D-RIBULOSE 1,5-BISPHOSPHATE,
 CC THE PRIMARY EVENT IN PHOTOSYNTHETIC CARBON DIOXIDE FIXATION, AS
 CC WELL AS THE OXIDATIVE FRAGMENTATION OF THE PENTOSE SUBSTRATE IN
 CC THE PHOTORESPIRATION PROCESS. BOTH REACTIONS OCCUR SIMULTANEOUSLY
 CC AND IN COMPETITION AT THE SAME ACTIVE SITE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
 CC 2 3-PHOSPHO-D-GLYCERATE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
 CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
 DR EMBL; X04975; CHMSRBCL.
 DR PIR; A25578; RKAALC.
 DR PROSITE; PS00157; RUBISCO_LARGE.
 KW PHOTOSYNTHESIS; CARBON DIOXIDE FIXATION; PHOTORESPIRATION;
 KW LYASE; OXIDOREDUCTASE; MONOOXYGENASE; CHLOROPLAST; ACETYLATION.
 FT PROPEP 1 2 BY SIMILARITY.
 FT CHAIN 3 474 RUBISCO LARGE CHAIN.
 FT MOD_RES 3 3 ACETYLATION (BY SIMILARITY).
 FT ACT_SITE 201 201 BINDING OF CO(2) ACTIVATES THE ENZYME.
 SQ SEQUENCE 474 AA; 52626 MW; 1066493 CN;

Initial Score = 8 Optimized Score = 8 Significance = 6.22
 Residue Identity = 50% Matches = 8 Mismatches = 8
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSMGTVFVGYGPTFK
 || || || ||
 TGTWTTVWTDGLTSLDRYKGRGYHIEPVAGEETQFIAYVAYPLDLFEEGSVNYMFTSIVGNVFGFKALRALR
 70 80 90 100 110 X 120 130
 LEDLRIPAAAYVKTFQGPPOGIQVERDKLNKYGRPLLGCITKPKL
 140 150 160 170

4. US-08-249-182-9 (1-16)

SODC_NPVAC PUTATIVE SUPEROXIDE DISMUTASE (CU-ZN) (EC 1.15.1.1)

ID SODC_NPVAC STANDARD; PRT; 151 AA.
 AC P24705;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE PUTATIVE SUPEROXIDE DISMUTASE (CU-ZN) (EC 1.15.1.1).
 GN SOD.
 OS AUTOGRAPHAL CALIFORNICA NUCLEAR POLYHEDROSIS VIRUS (ACMPV).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L1;
 RM 91335744
 RA TOMALSKI M.D., ELDRIDGE R., MILLER L.K.;

RL VIRUS: 184-149-181-1991).
 CC -!- FUNCTION: NONESSENTIAL FOR NORMAL VIRUS REPLICATION. COULD BE
 CC EITHER NON-FUNCTIONAL OR WITH A LOW ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
 CC -!- SIMILARITY: TO CU-ZN SUPEROXIDE DISMUTASES.
 DR EMBL; M68862; BANPASOD.
 DR PIR; A40564; DSNVAC.
 DR PROSITE; PS00087; SOD_CU_ZN_1.
 DR PROSITE; PS00332; SOD_CU_ZN_2.
 KW LATE PROTEIN; OXIDOREDUCTASE; COPPER; ZINC.
 FT METAL 43 43 COPPER (BY SIMILARITY).
 FT METAL 45 45 COPPER (BY SIMILARITY).
 FT METAL 60 60 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 68 68 ZINC (BY SIMILARITY).
 FT METAL 77 77 ZINC (BY SIMILARITY).
 FT METAL 80 80 ZINC (BY SIMILARITY).
 FT METAL 118 118 COPPER (BY SIMILARITY).
 FT DISULFID 54 144 BY SIMILARITY.
 SQ SEQUENCE 151 AA; 16182 MW; 113718 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

```

                                X      10      X
                                VNSMGTVFVGYGPTFK
                                || | | |||
    HEYGDTSNGCTSAGEHFNPTNEDHGAPDAEIRHVGD LGNIKSAGYNSLTEVNMMDNVMSLYGPHNIIGRSLV
      50      60      70      80      90      X 100      110

    VHTDKDDLGLTDHPLSKTTGNSGGRLGCGIIAICK
      120      130      140      150
  
```

5. US-08-249-182-9 (1-16)

UNG_FOWP1 URACIL-DNA GLYCOSYLASE (EC 3.2.2.-).

ID UNG_FOWP1 STANDARD; PRT; 218 AA.
 AC P21968;
 DT 01-AUG-1991 (REL. 19, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
 DE URACIL-DNA GLYCOSYLASE (EC 3.2.2.-).
 GN FPD4.
 OS FOWLPOX VIRUS (STRAIN FP-1).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC AVIPOXVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90324937
 RA TARTAGLIA J., WINSLOW J., GOEBEL S., JOHNSON G.P., TAYLOR J.,
 RA PAOLETTI E.;
 RL J. GEN. VIROL. 71:1517-1524(1990).
 CC -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE
 CC AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA
 CC POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE.
 CC -!- SIMILARITY: DISTANTLY, BUT SIGNIFICANTLY RELATED TO ALL OTHER
 CC SPECIES UNG.
 DR EMBL; X17202; POFPHIND.
 DR PIR; A35216; A35216.
 DR PROSITE; PS00130; U_DNA_GLYCOSYLASE.
 KW HYDROLASE; GLYCOSIDASE; DNA REPAIR.
 SQ SEQUENCE 218 AA; 25563 MW; 252913 CN;

Initial Score = 7 Optimized Score = 8 Significance = 5.18
 Residue Identity = 50% Matches = 8 Mismatches = 8

Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSMGTVFVGYGPTFK
  || || || |
NYYLSCREGEAKSHKIFWERLADVFINHIAAYVSVFYFLGKSDFSNFRSILNSPTTVVVGYPAAARNRQFDY
120      130      140      150      160      170      180      X 190

DETFEIVNTLLELKNEPRINWVQGFET
200      210

```

6. US-08-249-182-9 (1-16)

CD8A_RAT T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA CHAIN PRECUR

ID CD8A_RAT STANDARD; PRT; 236 AA.
AC P07725;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA CHAIN PRECURSOR (CD8 ANTIGEN 32
DE KD CHAIN) (OX-8 MEMBRANE ANTIGEN).
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RM 86030231
RA JOHNSON P., GAGNON J., BARCLAY A.N., WILLIAMS A.F.;
RL EMBO J. 4:2539-2545(1985).
CC -!- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T CELLS THAT INTERACT
CC WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN
CC THE PROCESS OF T CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO
CC CLASS MHC MOLECULES ALPHA-3 DOMAINS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SUBUNIT: IN GENERAL HETERODIMER OF AN ALPHA AND A BETA CHAIN
CC LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORMS HOMODIMERS.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
DR EMBL; X03015; RNANTOX8.
DR PIR; A24637; A24637.
KW IMMUNOGLOBULIN FOLD; TRANSMEMBRANE; T-CELL; ANTIGEN; GLYCOPROTEIN;
KW MHC; SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 236 T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA
FT CHAIN.
FT DOMAIN 27 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 212 POTENTIAL.
FT DOMAIN 213 236 CYTOPLASMIC (POTENTIAL).
FT DISULFID 47 119 POTENTIAL.
FT CARBOHYD 63 63 PROBABLE.
SQ SEQUENCE 236 AA; 26196 MW; 307073 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

```

X      10      X
VNSMGTVFVGYGPTFK
  || | | ||
DPNLFSARKENNKYILTLKSFSTKNQGYVFCSTNSVMYFSLPVVFQKVNSIITKPVTRAPTVPPTGT
90      100      110      120      130      140      150      X 160

PRPLRPEACRPGASGSVEGMGLGFACDIYIWAPLAGICAVLLLS
170      180      190      200

```

SC13_YEAST PROTEIN TRANSPORT PROTEIN SEC13.

ID SC13_YEAST STANDARD; PRT: 297 AA.
 AC Q04491;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE PROTEIN TRANSPORT PROTEIN SEC13.
 GN SEC13.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.
 RM 93163112
 RA PRYER N.K., SALAMA N.R., SCHEKMAN R., KAISER C.A.;
 RL J. CELL BIOL. 120:865-875(1993).
 CC -!- FUNCTION: REQUIRED IN VESICLE BIOGENESIS AT A STEP BEFORE OR
 CC CONCURRENT WITH THE RELEASE OF TRANSPORT VESICLES FROM THE ER
 CC MEMBRANE. REQUIRED FOR GERMINATION AND/OR GROWTH AT 24 DEGREE C.
 CC MIGHT INTERACT WITH PROTEINS SEC23 AND SAR1.
 CC -!- SUBUNIT: FORMS AN ACTIVE 700 KD LARGE COMPLEX WITH OTHER PROTEINS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. PERIPHERALLY ASSOCIATED WITH
 CC MEMBRANES.
 CC -!- SIMILARITY: SOME SIMILARITY TO THE BETA TRANSDUCIN FAMILY TRP-ASP
 CC DOMAINS.
 DR EMBL; L05929; SCSEC13P.
 DR PIR; S30803; S30803.
 DR PIR; A45442; A45442.
 KW TRANSPORT; PROTEIN TRANSPORT; MEMBRANE; ENDOPLASMIC RETICULUM;
 KW REPEAT.
 FT DOMAIN 8 281 6 X APPROXIMATE REPEATS.
 FT REPEAT 8 35 1.
 FT REPEAT 52 82 2.
 FT REPEAT 98 127 3.
 FT REPEAT 144 185 4.
 FT REPEAT 203 234 5.
 FT REPEAT 257 281 6.
 FT MUTAGEN 224 224 S->K: GROWTH INHIBITED ABOVE 30 C.
 FT MUTAGEN 262 262 W->R: GROWTH INHIBITED ABOVE 30 C.
 FT MUTAGEN 266 266 G->D: GROWTH INHIBITED ABOVE 34 C.
 SQ SEQUENCE 297 AA; 33043 MW; 480631 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSHQTVFVGYGPTFK
 ||| | |||

HEGPVWRVDWAHPKFGTILASCSYDGKVLWKEENGWRSQIAVHAVHSASVNSVQWAPHEYGPLLLVASSDG
 60 70 80 90 100 X 110 120

KVSVVEFKENGTTSPIIIDAHAGVNSASWAPATIEEDGEHNGT
 130 140 150 160

8. US-08-249-182-9 (1-16)

DHE4_UNKP NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4)

ID DHE4_UNKP STANDARD; PRT: 446 AA.
 AC P14657;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4) (NADP-GDH).

OS UNKNOWN PROKARYOTIC ORGANISM.
 OC PROKARYOTA; NOT YET CLASSIFIED.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 90098893
 RA COCK J.M., SCHMIDT R.R.;
 RL NUCLEIC ACIDS RES. 17:10500-10500(1989).
 CC -!- SEQUENCE ORIGINATES FROM AN ORGANISM CONTAMINATING A CHLORELLA
 CC SORDKINIANA CULTURE, PROBABLY A BACTERIUM.
 CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE + H(2)O + NADP(+) = 2-OXOGLUTARATE
 CC + NH(3) + NADPH.
 CC -!- SUBUNIT: HOMOHEXAMER.
 CC -!- SIMILARITY: TO OTHER GLUTAMATE DEHYDROGENASES (EC 1.4.1.2 AND
 CC EC 1.4.1.3), AND TO LEUCINE AND PHENYLALANINE DEHYDROGENASES.
 DR EMBL; X16399; XXGDH.
 DR PIR; S06938; S06938.
 DR PROSITE; PS00074; GLF_DEHYDROGENASE.
 KW OXIDOREDUCTASE; NADP.
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 446 AA; 48490 MW; 922941 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSMGTVFVGYGPTFK
 || || ||
 LERLVEPERIIQFRVSWVDRGQVQVNRAFRVQFNSAIGPYKGGMRFHPSVNLKFLGFQTFKNALTTL
 60 70 80 90 100 110 X 120
 PMGGGKGGSDFDPKGKSQGRIMRFCQALMTELYRHLGPD TDVPA
 130 140 150 160

9. US-08-249-182-9 (1-16)

DHE4_SALTY NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4)

ID DHE4_SALTY STANDARD; PRT; 447 AA.
 AC P15111;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4) (NADP-GDH).
 GN GDHA.
 OS SALMONELLA TYPHIMURIUM.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 89255551
 RA BANSAL A., DAYTON M.A., ZALKIN H., COLMAN R.F.;
 RL J. BIOL. CHEM. 264:9827-9835(1989).
 CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE + H(2)O + NADP(+) = 2-OXOGLUTARATE
 CC + NH(3) + NADPH.
 CC -!- SUBUNIT: HOMOHEXAMER.
 CC -!- SIMILARITY: TO OTHER GLUTAMATE DEHYDROGENASES (EC 1.4.1.2 AND
 CC EC 1.4.1.3), AND TO LEUCINE AND PHENYLALANINE DEHYDROGENASES.
 DR EMBL; M24021; STGDHA.
 DR PROSITE; PS00074; GLF_DEHYDROGENASE.
 KW OXIDOREDUCTASE; NADP.
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 447 AA; 48560 MW; 955116 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9

gaps - - - - - conservative substitutions - - -

X 10 X
VNSMGTVFVVGYPGPTFK
|| || |||

LERLVEPERVIQFRVVHLDDKNQVQVNRARVQFNSAIGPYKGGMRFHPSVNLKFLGFEGTFKNALTTL

60 70 80 90 100 110 X 120

PMGGGKGGSDFDPKGKSEGEVMRFCQALMTELYRHLGPDTDVPA

130 140 150 160

10. US-08-249-182-9 (1-16)

DHE4_ECOLI NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4)

ID DHE4_ECOLI STANDARD; PRT; 447 AA.
AC P00370;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE NADP-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.4) (NADP-GDH).
GN GDHA.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-36.
RC STRAIN=K12;
RM 83272967
RA MCPHERSON M.J., WOOTTON J.C.;
RL NUCLEIC ACIDS RES. 11:5257-5266(1983).
RN [2]
RP SEQUENCE FROM N.A.
RM 84209849
RA VALLE F., BECERRIL B., CHEN E., SEEBURG P.H., HEYNEKER H., BOLIVAR F.;
RL GENE 27:193-199(1984).
CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE + H(2)O + NADP(+) = 2-OXOGLUTARATE
CC + NH(3) + NADPH.
CC -!- SUBUNIT: HOMOHEXAMER.
CC -!- SIMILARITY: TO OTHER GLUTAMATE DEHYDROGENASES (EC 1.4.1.2 AND
CC EC 1.4.1.3), AND TO LEUCINE AND PHENYLALANINE DEHYDROGENASES.
DR EMBL; X00988; ECGDHA.
DR EMBL; K02499; ECGDHAK.
DR PIR; A00382; DEECEN.
DR PIR; A22413; A22413.
DR ECO2DBASE; G043.6; 5TH EDITION.
DR ECO2DBASE; G043.7; 5TH EDITION.
DR ECGENE; EG10372; GDHA.
DR PROSITE; PS00074; GLF_DEHYDROGENASE.
KW OXIDOREDUCTASE; NADP.
FT ACT_SITE 128 128 BY SIMILARITY.
FT CONFLICT 385 385 A -> P (IN REF. 2).
SQ SEQUENCE 447 AA; 48581 MW; 947292 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
Residue Identity = 43% Matches = 7 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

X 10 X
VNSMGTVFVVGYPGPTFK
|| || |||

LERLVEPERVIQFRVVWVDDRNIQVNRARVQFSSAIGPYKGGMRFHPSVNLKFLGFEGTFKNALTTL

60 70 80 90 100 110 X 120

PMGGGKGGSDFDPKGKSEGEVMRFCQALMTELYRHLGADTDVPA

130 140 150 160

11. US-08-249-182-9 (1-16)

AMPL_BOVIN CYTOSOL AMINOPEPTIDASE (EC 3.4.11.1) (LEUCINE AMIN

ID AMPL_BOVIN STANDARD; PRT: 478 AA.
AC P00727;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE CYTOSOL AMINOPEPTIDASE (EC 3.4.11.1) (LEUCINE AMINOPEPTIDASE) (LAP)
DE (LEUCYL AMINOPEPTIDASE) (PROLINE AMINOPEPTIDASE (EC 3.4.11.5) (PROLYL
DE AMINOPEPTIDASE)).
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE.
RC TISSUE=LENS;
RM 82213853
RA CUYPERS H.T., VAN LOON-KLAASSEN L.A.H., VREE EGBERTS W.T.M.,
RA DE JONG W.W., BLOEMENDAL H.;
RL J. BIOL. CHEM. 257:7077-7085(1982).
RN [2]
RP SEQUENCE.
RM 82213854
RA CUYPERS H.T., VAN LOON-KLAASSEN L.A.H., VREE EGBERTS W.T.M.,
RA DE JONG W.W., BLOEMENDAL H.;
RL J. BIOL. CHEM. 257:7086-7091(1982).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RM 90370887
RA BURLEY S.K., DAVID P.R., TAYLOR A., LIPSCOMB W.N.;
RL PROC. NATL. ACAD. SCI. U.S.A. 87:6878-6882(1990).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RM 92194311
RA BURLEY S.K., DAVID P.R., SWEET R.M., TAYLOR A., LIPSCOMB W.N.;
RL J. MOL. BIOL. 224:113-140(1992).
CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
CC TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
CC UNSUBSTITUTED AMINO-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES.
CC -!- CATALYTIC ACTIVITY: AMINOACYL-PEPTIDE + H(2)O = AMINO ACID +
CC PEPTIDE.
CC -!- COFACTOR: BINDS TWO ZINC IONS PER SUBUNIT. ONE ZINC ION IS TIGHTLY
CC BOUND AND ESSENTIAL FOR ENZYME ACTIVITY, WHILE THE SECOND METAL
CC COORDINATION SITE CAN BE OCCUPIED BY ZINC, MAGNESIUM OR MANGANESE
CC TO GIVE ENZYMES OF DIFFERENT ACTIVITIES.
CC -!- ENZYME REGULATION: INHIBITED BY BESTATIN.
CC -!- SUBUNIT: HOMOHEXAMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE CYTOSOL AMINOPEPTIDASE FAMILY.
DR PIR: A00907; APBOL.
DR PDB: 1LAP; 15-OCT-91.
DR PDB: 1BPM; 15-JUL-93.
DR PDB: 1BPN; 15-JUL-93.
DR PROSITE: PS00631; CYTOSOL_AP.
KW HYDROLASE; AMINOPEPTIDASE; ACETYLATION; ZINC; 3D-STRUCTURE; MAGNESIUM;
KW MANGANESE.
FT MOD_RES 1 1 ACETYLATION.
FT METAL 250 250 ZINC (2).
FT METAL 255 255 ZINC (1 AND 2).
FT METAL 273 273 ZINC (2).
FT METAL 332 332 ZINC (1).
FT METAL 334 334 ZINC (1 AND 2).
FT ACT_SITE 262 262 POTENTIAL.

FT STRAND 465 466
 FT TURN 469 470
 FT HELIX 471 472
 SQ SEQUENCE 478 AA; 51692 MW; 1107347 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSM@TVFVGYGPTFK
 |||| | ||
 NLKSASIKTDVFI RPKSWIEE@EMGSFLSVAKGSEPPVFLEIHYKGSNPASEPPLVFGKGITFDSGGISI
 190 200 210 220 230 240 250 X 260
 KAAANMDLMRADMGGAATICS AIVSAAKLDPINIVGLAPLCEN
 270 280 290 300

12. US-08-249-182-9 (1-16)

HEMA_PI3B HEMAGGLUTININ-NEURAMINIDASE (EC 3.2.1.18).

ID HEMA_PI3B STANDARD; PRT; 572 AA.
 AC P06167;
 DT 01-JAN-1988 (REL. 06, CREATED)
 DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 DE HEMAGGLUTININ-NEURAMINIDASE (EC 3.2.1.18).
 GN HN.
 OS BOVINE PARAINFLUENZA 3 VIRUS.
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; PARAMYXOVIRIDAE;
 OC PARAMYXOVIRUSES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 87174819
 RA SUZU S., SAKAI Y., SHIODA T., SHIBUTA H.;
 RL NUCLEIC ACIDS RES. 15:2945-2958(1987).
 CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 CC TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEUROAMINIDASE
 CC ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 CC THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 CC GLYCOPROTEINS.
 CC -!- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 CC N-TERMINAL HYDROPHOBIC SEQUENCE.
 CC -!- SIMILARITY: TO HEMAGGLUTININ-NEURAMINIDASES FROM OTHER
 CC PARAMYXOVIRUSES.
 DR EMBL; Y00114; PAMBPIV3.
 DR PIR; B27218; HNNZB3.
 KW HYDROLASE; HEMAGGLUTININ; ENVELOPE PROTEIN; GLYCOPROTEIN;
 KW TRANSMEMBRANE.
 FT DOMAIN 1 30 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 31 53 POTENTIAL.
 FT DOMAIN 54 572 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 308 308 POTENTIAL.
 FT CARBOHYD 351 351 POTENTIAL.
 FT CARBOHYD 448 448 POTENTIAL.
 FT CARBOHYD 523 523 POTENTIAL.
 FT CARBOHYD 570 570 POTENTIAL.
 SQ SEQUENCE 572 AA; 64590 MW; 1764491 CN;

Initial Score = 7 Optimized Score = 7 Significance = 5.18
 Residue Identity = 43% Matches = 7 Mismatches = 9
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 VNSM@TVFVGYGPTFK

FT	ACT_SITE	336	336
FT	STRAND	3	8
FT	STRAND	20	20
FT	HELIX	22	30
FT	TURN	31	33
FT	HELIX	34	42
FT	TURN	48	49
FT	STRAND	51	58
FT	TURN	59	60
FT	STRAND	61	68
FT	TURN	72	73
FT	STRAND	76	77
FT	TURN	78	81
FT	STRAND	82	83
FT	HELIX	84	102
FT	TURN	103	104
FT	STRAND	107	110
FT	TURN	113	114
FT	HELIX	116	126
FT	TURN	127	127
FT	TURN	132	133
FT	STRAND	142	145
FT	TURN	149	150
FT	HELIX	151	172
FT	TURN	175	177
FT	HELIX	180	194
FT	STRAND	198	203
FT	HELIX	205	210
FT	TURN	211	212
FT	HELIX	214	220
FT	TURN	221	222
FT	STRAND	228	235
FT	TURN	240	241
FT	STRAND	245	249
FT	STRAND	251	255
FT	TURN	258	259
FT	TURN	265	266
FT	HELIX	267	272
FT	TURN	273	274
FT	HELIX	275	289
FT	TURN	290	291
FT	STRAND	295	305
FT	TURN	309	310
FT	TURN	314	315
FT	STRAND	316	319
FT	TURN	321	322
FT	STRAND	325	328
FT	TURN	331	332
FT	HELIX	334	348
FT	TURN	349	349
FT	STRAND	353	358
FT	HELIX	362	368
FT	TURN	369	370
FT	STRAND	373	377
FT	HELIX	380	392
FT	STRAND	396	398
FT	HELIX	403	409
FT	TURN	413	414
FT	STRAND	417	418
FT	TURN	425	426
FT	HELIX	427	436
FT	TURN	437	438
FT	STRAND	444	448
FT	HELIX	450	452
FT	STRAND	454	455
FT	TURN	460	461

POTENTIAL.